QUANTIFICATION OF THE EFFECTS OF DATA DENIAL

AND LIMITATION IN MM5 INITIALIZATION ON

FORECAST ACCURACY

THESIS

Robert W. Evans, First Lieutenant, USAF

AFIT/GM/ENP/02M-03

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

Report Documentation Page

Report Date 26 Mar 02	**Report Type** Final	**Dates Covered (from... to)** Jun 01 - Mar 02

Title and Subtitle Quantification of the Effects of Data Denial and Limitation in MM5 Initialization on Forecast Accuracy	**Contract Number**
	Grant Number
	Program Element Number

Author(s) 1t Lt Robert W. Evans, USAF	**Project Number**
	Task Number
	Work Unit Number

Performing Organization Name(s) and Address(es) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 P Street, Bldg 640 WPAFB, OH 45433-7765	**Performing Organization Report Number** AFIT/GM/ENP/02-03

Sponsoring/Monitoring Agency Name(s) and Address(es) AFWA/DNXM ATTN: Jerry Wegiel 106 Peacekeeper Dr., Offutt AFB NE 68113-4039 88WS/DO ATTN: Maj Peter Roohr 2049 Monahan Way, Bldg 91 WPAFB OH 45433-7204	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
The original document contains color images.

Abstract

Using the 3-Dimensional Variational Analysis data assimilation scheme and the Pennsylvania State University (PSU)/National Center for Atmospheric Research (NCAR) Mesoscale Model 5 (MM5) as configured and run operationally by the Air Force Weather Agency (AFWA) in the European theater, input observations were denied in three different categories: total, upper air, and surface observation denial. Two control groups were run using all available data as received by AFWA. The main control group used a 6-hour old first guess as a baseline. The data denied test cases and the secondary control group used a 30-hour old first guess because it was not possible to deny data from the first guess. The secondary control group was used to estimate errors resulting from the use of different first guess forecasts between the main control group and the test cases for all forecast times. The analyses show statistically significant differences between the main control group and test cases in almost every instance. However, the surface comparisons provide little evidence of significant meteorological differences due to the relatively small magnitude of the differences in root mean square error and bias. In the upper air analyses, the largest differences emerged at the model initialization time. The magnitude of the differences rapidly diminishes as the solution of the test cases converge toward that of the main control group throughout the forecast period. Errors associated with the use of different initial and boundary conditions account for most of the resulting differences.

Subject Terms

Mesoscale Model, MM5, 3DVAR, 3-Dimensional Variational Analysis, Air Force Weather Agency, AFWA MM5, Numerical Weather Prediction, Model Coparison, Atmospheric Temperature, Weather

Report Classification unclassified	Classification of this page unclassified
Classification of Abstract unclassified	Limitation of Abstract UU

Number of Pages

120

AFIT/GM/ENP/02M-03

QUANTIFICATION OF THE EFFECTS OF DATA DENIAL AND LIMITATION IN

MM5 INITIALIZATION ON FORECAST ACCURACY

THESIS

Presented to the Faculty

Department of Engineering Physics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Meteorology

Robert W. Evans, B.S.

First Lieutenant, USAF

March 2002

QUANTIFICATION OF THE EFFECTS OF DATA DENIAL AND LIMITATION IN

MM5 INITIALIZATION ON FORECAST ACCURACY

Robert W. Evans, B.S.

First Lieutenant, USAF

Approved:

_____
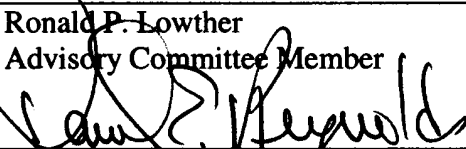Michael K. Walters
Advisory Committee Chairman

7 MAR 12
_____
Date

_____
Ronald P. Lowther
Advisory Committee Member

7 MAR 02
_____
Date

_____
Daniel E. Reynolds
Advisory Committee Member

7 Mar 02
_____
Date

Table of Contents

Page

List of Figures

viii

List of Tables

Abstract

Denied and limited model input data degrade the accuracy of numerical weather prediction (NWP) model forecasts. Inaccurate forecasts can negatively impact all military operations and public safety in general.

Using the 3-Dimensional Variational Analysis data assimilation scheme and the Pennsylvania State University (PSU)/National Center for Atmospheric Research (NCAR) Mesoscale Model 5 (MM5) as configured and run operationally by the Air Force Weather Agency (AFWA) in the European theater, input observations were denied in three different categories: total, upper air, and surface observation denial. Two control groups were run using all available data as received by AFWA. The main control group used a 6-hour old first guess as a baseline. The data denied test cases and the secondary control group used a 30-hour old first guess because it was not possible to deny data from the first guess. The secondary control group was used to estimate errors resulting from the use of different first guess forecasts between the main control group and the test cases for all forecast times.

The analyses show statistically significant differences between the main control group and test cases in almost every instance. However, the surface comparisons provide little evidence of significant meteorological differences due to the relatively small magnitude of the differences in root mean square error and bias. In the upper air analyses, the largest differences emerged at the model initialization time. The magnitude of the differences rapidly diminishes as the solution of the test cases converge toward that of the main control group throughout the forecast period. Errors associated with the use of different initial and boundary conditions account for most of the resulting differences.

QUANTIFICATION OF THE EFFECTS OF DATA DENIAL AND

LIMITATION IN MM5 INITIALIZATION ON FORECAST ACCURACY

I. Introduction

1.1. Motivation

Numerical weather prediction (NWP) models are complex systems of mathematical equations written to simulate atmospheric motions taking place at and above the earth's surface. The ability to develop and run more complex models has exploded along with the development of faster and more powerful computer processors allowing more accurate calculations of air flows and more complete representations of factors that affect meteorological changes on a smaller scale (NOAA, 2001). These advances in computer technology have made the use of limited area, or mesoscale numerical models increasingly more attractive to the government, commercial, and educational institutions for both research and operational applications (Stenger, 2000).

Numerical models are only as good as the information that initializes them. All models use some sort of data assimilation scheme, which allows information about current and future states of the atmosphere to be ingested. In general, current observational data are compiled from land and sea surface measurements, radiosondes, pilot reports, aircraft, and satellites. Data relative to future states of the atmosphere are ingested from earlier model runs from the same model or from a different model entirely. In mesoscale models, such as

the Pennsylvania State University (PSU)/National Center for Atmospheric Research (NCAR) Mesoscale Model 5 (MM5), data assimilation is accomplished through the comparison of observed data and forecasts from larger, global scale models such as the Aviation Global Model (AVN) or the Navy Operational Global Atmospheric Prediction System (NOGAPS) model. The data assimilation process is explained in greater detail in Chapter II.

The Air Force Weather Agency (AFWA) at Offutt AFB, NE has proposed this topic to enhance their understanding of the limitations of meteorological models due to denied or limited input data. AFWA runs the MM5 operationally with windows set up over specific regions, or Theaters of Operation (Ritz et al., 2001). AFWA currently uses the Mesoscale Data Assimilation System/Multivariate Optimal Interpolation (MDAS/MVOI) data assimilation scheme. However, they are in the testing phase of converting from the MDAS/MVOI to the Three-Dimensional Variational (3DVAR) data analysis assimilation scheme. They plan to have the transition complete by mid-to-late 2002 (Wegiel, 2001). This thesis will seek to quantify the effects of denied and limited data on the forecast accuracy of the MM5 as run by AFWA using the 3DVAR scheme.

In general, the more usable observations that are ingested at the time of model initialization, the more accurate the model's forecast will ultimately be. However, there are three principal limitations associated with observational inputs. First, there is an inherent inhomogeneous distribution of data. This erratic distribution leads to clusters of observations in some regions with few or no observations in other regions. Another limitation is that the data are not always located in sensitive areas. Finally, all observations have errors, some of which are associated with the inherent limitations of the equipment, while others occur due to equipment malfunction. These limitations and errors become inflated as the model gets

further away from the initialization time. The 3DVAR data assimilation scheme is designed to alleviate some of this error by using previous forecasts to fill in gaps where observations are limited, however, there is still a need for more and improved observations.

Occasionally, some or all of the observation data do not reach AFWA in time to be assimilated into the newest forecast. In this situation, the MM5 is run with the most current information available. In the event of limited data, one would expect the output of NWP models to be degraded. However, limitation or denial of observed weather data could be the result of a number of factors. These include, but are not limited to, observation equipment failure or malfunction, communication problems, computer errors, or simply denial of data by an adversary. Over an extended period of time, this degradation could render the model's forecast virtually useless in certain regions to weather forecasters around the world. The effects of data denial could be anywhere from minimal, if the affected area is relatively small, to extreme in the case of total data denial over a large area.

Numerical models used operationally must run in a timely manner or their usefulness to forecasters is very limited. AFWA must run the MM5 within a limited amount of time after the initialization time in order for all of the computations to be completed. The output must then be made available to the operational meteorologists so it can be used as a forecasting tool. Due to these time sensitivities of the data and model run, it is not feasible for AFWA, or any other agency, to wait for all observations to be available for the start of the model initialization.

1.2. Problem Statement

Numerical weather prediction model forecasts are degraded when observations are unavailable at the initialization time. The degradation of the forecast can have severe adverse impacts on mission success and pose risks on personnel and resources. The magnitude of the degradation related to input data denial and limitation is unknown at this time with respect to 3DVAR used in conjunction with the MM5 model at AFWA.

1.3. Research Objectives

The goal of this thesis is to quantify the effects of data denial and limitation on the accuracy of MM5 model forecasts as run by AFWA using the 3DVAR data assimilation scheme over the European theater of operation. The European theater was chosen since it is the location for which the 3DVAR code was initially set up at AFWA. The following are specific objectives used in this study to arrive at a sound assessment of these effects:

1. Collect adequate observational and AVN model output data for the European theater in order to provide a statistically significant sample. These observational data include all of the previously mentioned observations from surface measurements, radiosondes, pilot reports, aircraft, and satellite sources. Surface observations are collected from 90 weather stations, and radiosonde observation (RAOB) data are collected from 68 stations releasing weather balloons at least twice daily.

2. Obtain necessary computing resources in addition to pre- and post-processing code in order to duplicate, to the greatest extent possible, the MM5 output being produced by AFWA. Exceptions to this are made so that the data assimilation will occur through

4

the use of the most recent version of 3DVAR available to AFWA, which is currently

in the developmental phase.

3. Run the MM5 model four times at each initialization time: once using all available

    observations (control run) and three times using different data denial configurations

    (test cases). Each set of four model outputs will cover a period of 24 hours with

    output generated at the initialization time and every three hours after that out to the

    24-hour forecast. The first run will use all available observation inputs using a 6-hour

    old AVN forecast to provide the first guess for the initial and boundaries conditions.

    All subsequent runs will be conducted with various limited observational inputs and,

    as a minimum, 30-hour old AVN forecasts as the first guess. Use of the 30-hour old

    first guess forecasts is necessary since it is beyond the scope of this research to also

    deny the observation inputs to the AVN model initialization process in order to

    duplicate a real world data denial situation. The method of model initialization will

    be discussed in detail in Chapter III.

4. Compute residuals of the difference between the output of the MM5 model initialized

    with all available input data and output from the runs initialized with limited input

    data. The root mean square error (RMSE) and bias of both model forecast outputs is

    then compared. The specific surface variables to be examined are sea level pressure,

    temperature, and winds. In the vertical, geopotential heights and vertical temperature

    profiles are compared. A thorough statistical analysis is used to quantify the effects

    the data denial and limitation had on the model output.

5. An additional control run will be made using the 30-hour old first guess forecast and

    all available observations. The new control run will be compared to the first one in

an effort to account for differences in the initial and boundary conditions between the primary control group and the data denied test cases. Comparisons between the two control runs will be made in the same manner as described previously.

Each of these tasks is approached with the idea of duplicating the AFWA production MM5 as much as possible. The primary exception to this duplication was in the use of the 3DVAR data assimilation scheme versus the MDAS/MVOI used by AFWA. For this reason, there cannot be a direct comparison between the current MM5 production output generated by AFWA and that generated by this research. Nevertheless, it was important to duplicate AFWA's production MM5 as much as possible since they plan to implement the use of the 3DVAR scheme in the future. A detailed comparison of the differences between the 3DVAR and MDAS/MVOI data assimilation schemes is found in Chapter II.

1.4. Research Importance

As the speed, accuracy, and resolution of NWP models increases, so does the dependence on their output as an aid to weather forecasting. In the Air Force, other factors such as reduced manning, high operations tempo, and forecaster inexperience also lead to an increasing dependence on model output as an aid in forecasting. Inaccurate forecasts of any kind present potential negative impacts on all military operations and public safety in general. Advance knowledge of favorable or adverse weather conditions is also critically important to military planners and strategists. Accurate weather forecasts give planners the information to make the most effective and efficient use of the assets at their disposal.

1.5. Thesis Organization

　　This chapter provided the motivation for this research followed by an overview of the problem, the goal/objectives of the research and its importance to military operations. Chapter II provides a detailed background on the MM5 and 3DVAR scheme. Chapter III discusses the methodology used to ensure valid and statistically sound comparisons were made. Chapter IV summarizes the results of the data collection and analysis process, and Chapter V presents conclusions and recommendations for future research based on the data and results found.

II. Background and Literature Review

2.1. Overview

This chapter presents the subject background necessary to understand the significance data denial or limitation has on the output of numerical weather prediction (NWP) models, specifically, the MM5 using the 3DVAR data assimilation scheme. First, a brief history of the MM5 at AFWA is provided. Then, Section 2.3 provides a synopsis of the MM5 and how it works with current data assimilation schemes. Finally, Section 2.4 provides a discussion of data assimilation in general and a more detailed discussion of the 3DVAR scheme.

2.2. History of the MM5 at AFWA

The Air Force Weather Agency (AFWA) runs the MM5 operationally and provides fine-scale, limited-area forecasts for military and civilian forecasters worldwide (Ritz et al., 2001). AFWA has run this particular model continually since 1997. From that time until January 2001, initialization of the model was accomplished by using the analysis portion of the Local Analysis and Prediction System (LAPS) from the Forecast Systems Laboratory. The version of LAPS used by AFWA was a successive correction scheme using a limited number of available observation types. In January 2001, LAPS was replaced by the Mesoscale Data Assimilation System/Multivariate Optimal Interpolation (MDAS/MVOI).

Since the MDAS/MVOI system had been used previously by AFWA in running the Relocatable Window Model, a major advantage of using it with the MM5 model was its ease of relocation to any portion of the globe. Another advantage was the system's ability to produce analyses using observational data varying greatly in quantity and quality from one

location to another. The MDAS/MVOI system currently used provides rapid, accurate

initialization fields for the MM5 model. Speed and accuracy are critical to the support of

worldwide operations (Ritz et al., 2001).

In June 2001, AFWA started the transition to the Three-Dimensional Variational

(3DVAR) data assimilation scheme. Currently 3DVAR has been successfully integrated

with the MM5 model in the European Theater (Wegiel, 2001). Its operation is being verified

by running both the 3DVAR and MDAS/MVOI data assimilation schemes simultaneously.

Full global integration of the 3DVAR data assimilation is expected to be complete by mid-to-

late 2002.

The primary purpose for upgrading to the 3DVAR data assimilation is to improve

model forecasts relative to those integrated from analyses produced from currently used

assimilation schemes (NCAR, 2001). In addition, the upgrade should enable a more

computationally efficient system and provide ease of integration into open architectures.

2.3. Overview of the MM5

This section provides a general description of the MM5 modeling system as run in

production by AFWA. For a more detailed description of the MM5, the reader is referred to

Dudhia et al., 2001.

2.3.1. Background. The MM5 is a mesoscale model initially developed in 1971 at

PSU in cooperation with NCAR. Both NCAR and PSU have worked extensively to improve

the model and it has evolved to a fifth-generation. The AFWA implementation of the MM5

consists of four preprocessors (TERRAIN, REGRID, MDAS/MVOI and INTERP), the main

model code (MM5 (V3.4)), and a post-processing routine. Figure 1 is a schematic diagram

representing the flow of data from the preprocessing modules through the post-processor.

For the current research, 3DVAR was substituted for MDAS/MVOI and the

Read/Interpolate/Plot (RIP) software was used for post-processing.

Input   **Programs**

USGS Terrain, Vegetation and Soil ———▶ [TERRAIN]
Data

AVN or NOGAPS Forecasts ——————▶ [REGRID]

Observations ——————————————▶ [MDAS/MVOI]

[INTERPF]

[MM5]

[GRAPH/RIP]
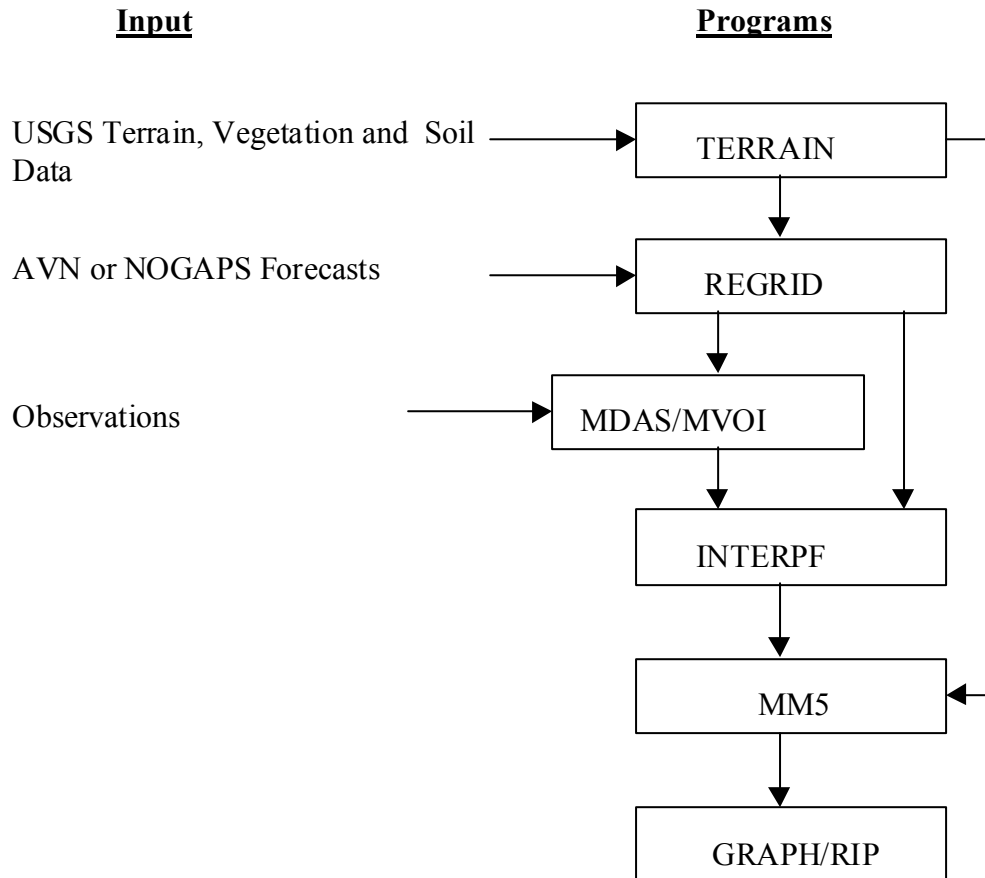
Figure 1. MM5 flow chart. Schematic diagram depicting the flow of data throughout the MM5 modeling
system (adapted from Dudhia et al., 2001).

TERRAIN is the first program used in the MM5 modeling system. TERRAIN

performs two primary tasks. First, it horizontally interpolates the regular latitude-longitude

terrain elevation and vegetation data onto the chosen mesoscale domains. TERRAIN's

second task is to produce the terrestrial data files for all of the mesoscale domains used later

by REGRID. In addition, TERRAIN calculates the constant fields of latitude and longitude,

map scale factors, and the Coriolis parameter which are required by the modeling system (Dudhia et al., 2001).

The second preprocessing program within the MM5 modeling system is REGRID. The purpose of this program is to read archived meteorological data on given pressure levels and interpolate them to the grid and map projection defined by TERRAIN. REGRID only performs two-dimensional interpolation for pressure levels and surface analyses (Dudhia et al., 2001). As seen in Figure 1, REGRID creates output for MDAS/MVOI or INTERPF depending on the input data. These output files are used as the first guess for an objective analysis if used by MVOI or as analyses to be directly interpolated into initial and boundary conditions by INTERPF. The data ingest process is somewhat different when using 3DVAR and will be explained in a later section.

The next preprocessor is the MDAS/MVOI data assimilation scheme, which is followed by INTERPF. INTERPF performs the data transformation necessary to go from the analysis programs to the model. These transformations include vertical interpolation, diagnostic computation and data reformatting. This program outputs three files used by MM5: model initial conditions, lateral boundary conditions, and the lower boundary conditions (Dudhia et al., 2001). After the MM5 has been run, the output data must be post-processed into a user-readable format. This is the function of the GRAPH or RIP programs.

2.3.2. Model Characteristics. As run by AFWA, the MM5 produces forecasts valid at 3-hour intervals out to 72 hours four times per day with initial analyses valid at 00, 06, 12 and 18Z. The MM5 allows for a four-dimensional data-assimilation (FDDA) and uses numerous physical parameterization schemes. The FDDA process permits data to be assimilated over time as well as in the three dimensions of space. When available, AFWA

uses the Aviation (AVN) model for initial and boundary conditions, which is explained in

section 2.4. Otherwise, the Navy Operational Global Analysis and Prediction System

(NOGAPS) data are used. The AVN is the preferred global scale model because its output

resolution is 1° X 1° versus that of NOGAPS, which is 2.5° X 2.5°. In addition, the AVN is

run four times per day compared to the two daily runs of NOGAPS.

 2.3.3. Vertical Coordinate System. The MM5 model uses sigma levels to define its

vertical coordinate system. Sigma (σ) is a dimensionless quantity that varies from zero to one

and is given by the following relationship:

$$\sigma = (p - p_t)/(p_s - p_t), \tag{1}$$

where p is a reference pressure at a given level, p_t is a specified constant top pressure and p_s

is the constant surface pressure. The reference pressure is a function of the constant surface

pressure, which only varies in two-dimensional space, not time (Dudhia et al., 2001). Figure

2 shows an example of a vertical cross section with 16 sigma levels (*K*). The lowest level is

represented where $\sigma = 1.0$ and is terrain-following. This is the bottom boundary for the

model. Sigma levels have the distinct advantage of being terrain following (Grell et al.

1994). Physical processes along the bottom boundary can be prescribed using physical

process parameterizations, as described later. The highest level, where $\sigma = 0.0$, is quasi-

horizontal and represents the top boundary for the model. The top layer acts as a physical

boundary limiting the flux of atmospheric properties across it and can also be parameterized

within the model physics (Pielke, 1984). As run by AFWA, the MM5 produces forecasts on

42 different full-sigma levels and 41 half-sigma levels (Ritz, 2001). Vertical velocity (*w*) is
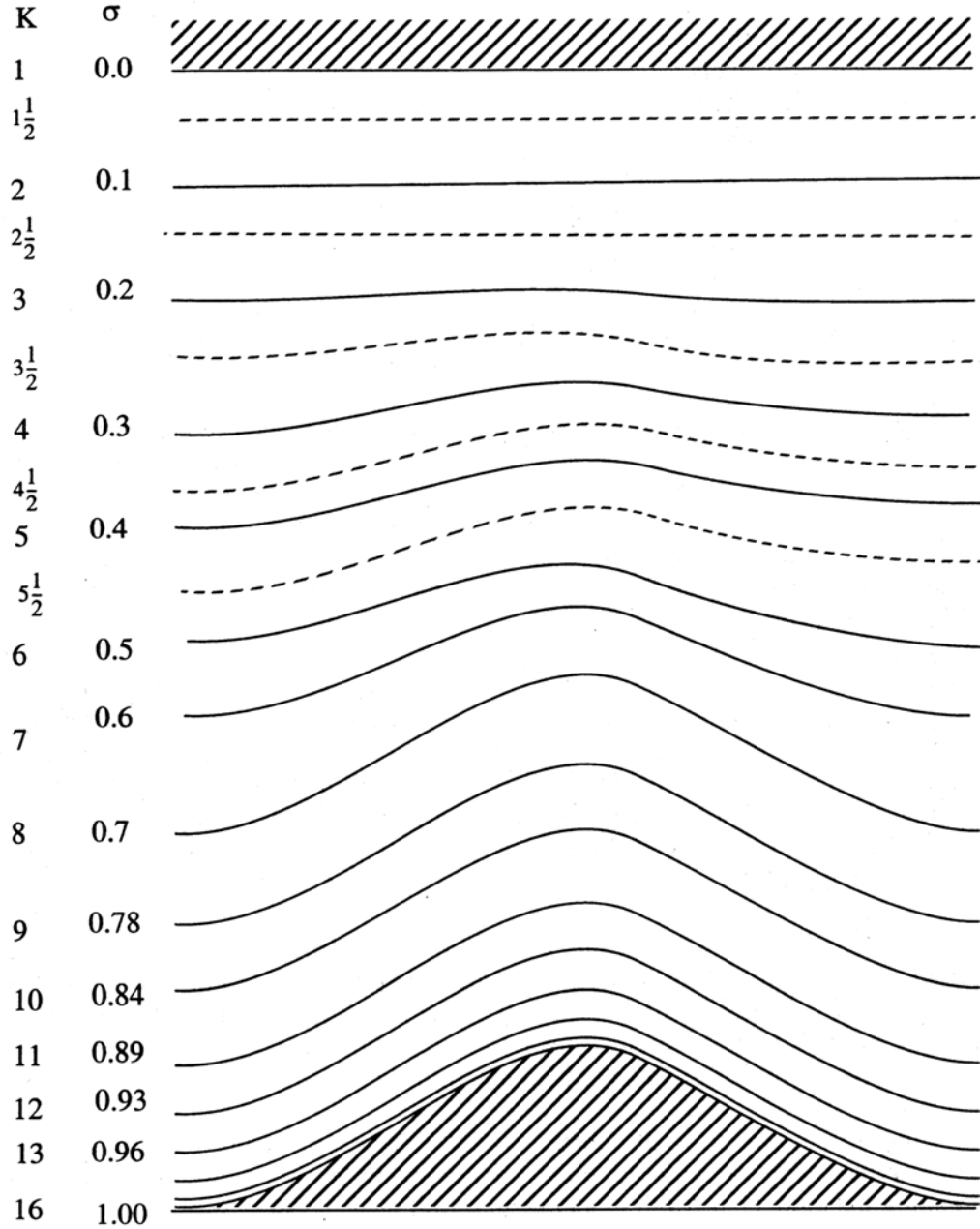
Figure 2. Sigma coordinates. Schematic representation of the vertical structure of the model. This example is for 15 vertical layers (or half-sigma levels). Dashed lines denote half-sigma levels, solid lines denote full-sigma levels (adapted from Dudhia et al., 2001).

the only parameter computed at full-sigma levels. The remaining parameters are all

calculated on the half-sigma levels (Dudhia et al., 2001).

Each sigma level is laid out in a grid consisting of cross (X) and dot (•) points. The MM5 uses an Arakawa-Lamb B Grid as seen in Figure 3. Both the eastward velocity component (u) and the northward component (v) of the wind are collocated at the corners, or dot points. Similarly, all remaining forecast variables are defined at the center of the grid squares, which are denoted by the cross points.
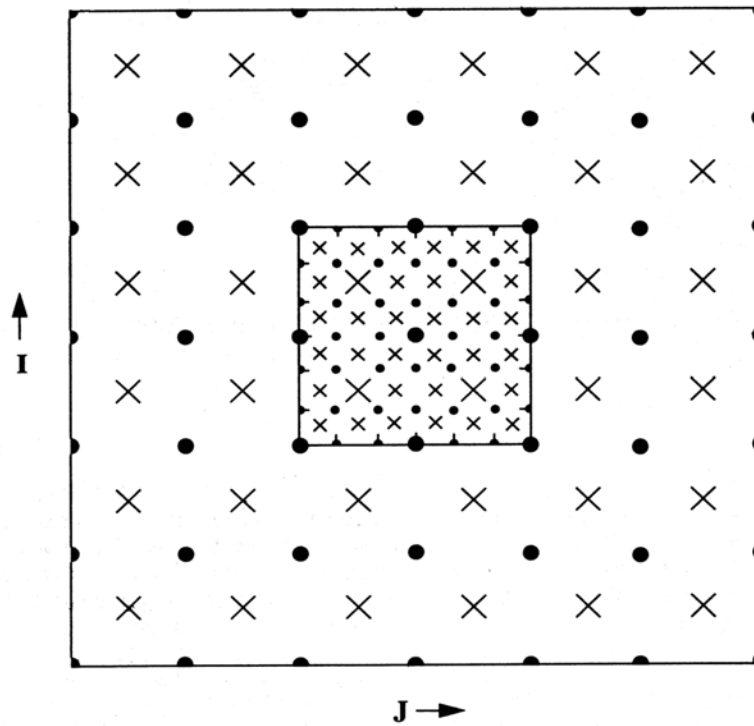
Figure 3. Arakawa B-Grid. Schematic diagram representing the horizontal Arakawa B-Grid staggering of the dot (•) and cross (X) points. The smaller inner box is a representative mesh staggering for a 3:1 coarse-grid distance to fine-grid distance ratio (adapted from Dudhia et al., 2001).

2.3.4. Finite Differencing. The MM5 uses second-order centered finite differences to represent all of the spatial gradients except for the precipitation fall term, which uses a first-order upstream scheme. Horizontal averaging is required to determine the gradient of the computed parameters on each sigma level. Vertical interpolations are performed to account for the variable vertical grid size (Dudhia et al., 2001).

Time differencing is accomplished using a second-order leapfrog scheme. However, because there are fast moving components (e.g. sound waves) that must be calculated on a shorter time step for stability, a time-splitting scheme is used. In the leapfrog scheme, detailed in Figure 4, the tendencies at time *n* are used to step the variables from time n-1 to n+1. The leapfrog scheme is used for the larger-scale forcing terms such as advection and Coriolis. A forward step is used for diffusion and microphysics where the tendencies are computed at time n-1. The forward step is used to step the variables from n-1 to n+1. The longer time step is too long for stability for fast moving waves such as sound waves, so the shorter time step depicted in the bottom portion of Figure 4 must be used instead. The short time step allows the variables to be updated more frequently (Dudhia et al., 2001).

Long (Leapfrog) step used for T, qv, qc, advection, physics, boundary, coriolis, and diffusion terms
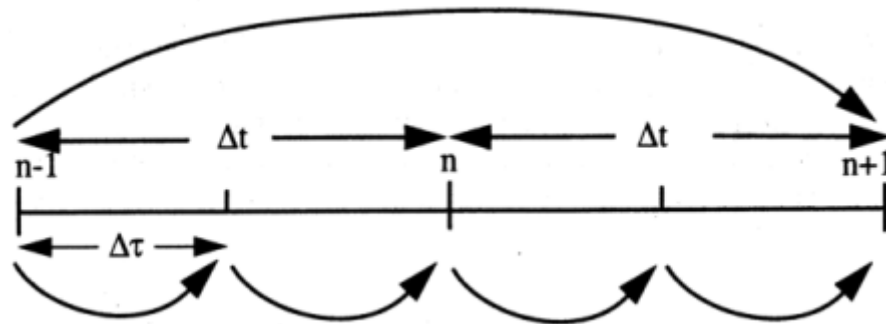
Short (Forward) step used for u, v, w and p' advanced (pressure gradients and divergence)

Figure 4. Leapfrog and time-splitting schemes. Schematic diagram of the leapfrog and time splitting scheme used for temporal finite differencing in the MM5 (adapted from Dudhia et al., 2001).

2.3.5. Parameterization Schemes. Atmospheric phenomena such as momentum, advection, vorticity, and divergence occur on a scale larger than that of the model grid resolution and are resolvable by the model. However, parameters such as moisture and heat

15

fluxes, planetary boundary layer processes, cumuliform cloud formation, and precipitation physics occur on a smaller scale than that of the grid and are not readily resolved by the model. Therefore, it is necessary to incorporate different parameterization schemes to represent these phenomena. Parameterization schemes provide the necessary interaction between grid scale processes and each other in order to compute parameters occurring on sub-grid scales.

2.4. Data Assimilation

Data assimilation is the process in which observations are ingested in order to provide corrections to the first guess forecast to describe the initial conditions for the next model run (COMET, 2001). This section describes the 3DVAR data assimilation process in more detail. Finally, there is a discussion of the advantages of 3DVAR over MDAS/MVOI.

2.4.1. Background. Observations are not specifically used to make a model analysis. They are, however, used to make small corrections to a short-range forecast used for initialization. This short-range forecast is assumed to be representative of actual conditions (COMET, 2001). An example of this is AFWA's use of the 6-hour AVN forecast from the 18Z run to initialize 00Z MM5 run. In a similar manner, the AVN forecast uses a previous AVN run with observations near the initialization time for its analysis. In this way, short-range forecasts retain information from older observations and provide background information into future analyses. It is important to note that the production MM5 as run by AFWA is never initialized with a previous MM5 model run.

The key to the initialization process is the assumption that the short-range forecast used is representative of actual conditions. This assumption allows the observations to be

used to make a series of small corrections that could improve an already good forecast rather than analyzing the observations directly and starting over each time. This assumption also provides a means of quality control; large discrepancies between the observations and the forecast are used to determine if the data are suspect or erroneous. Conversely, if the assumption is wrong and the forecast is bad, the subsequent forecast could also be bad. This cycle will continue until enough data are available to force changes in the model output bringing it back to values representing the actual atmospheric conditions (COMET, 2001).
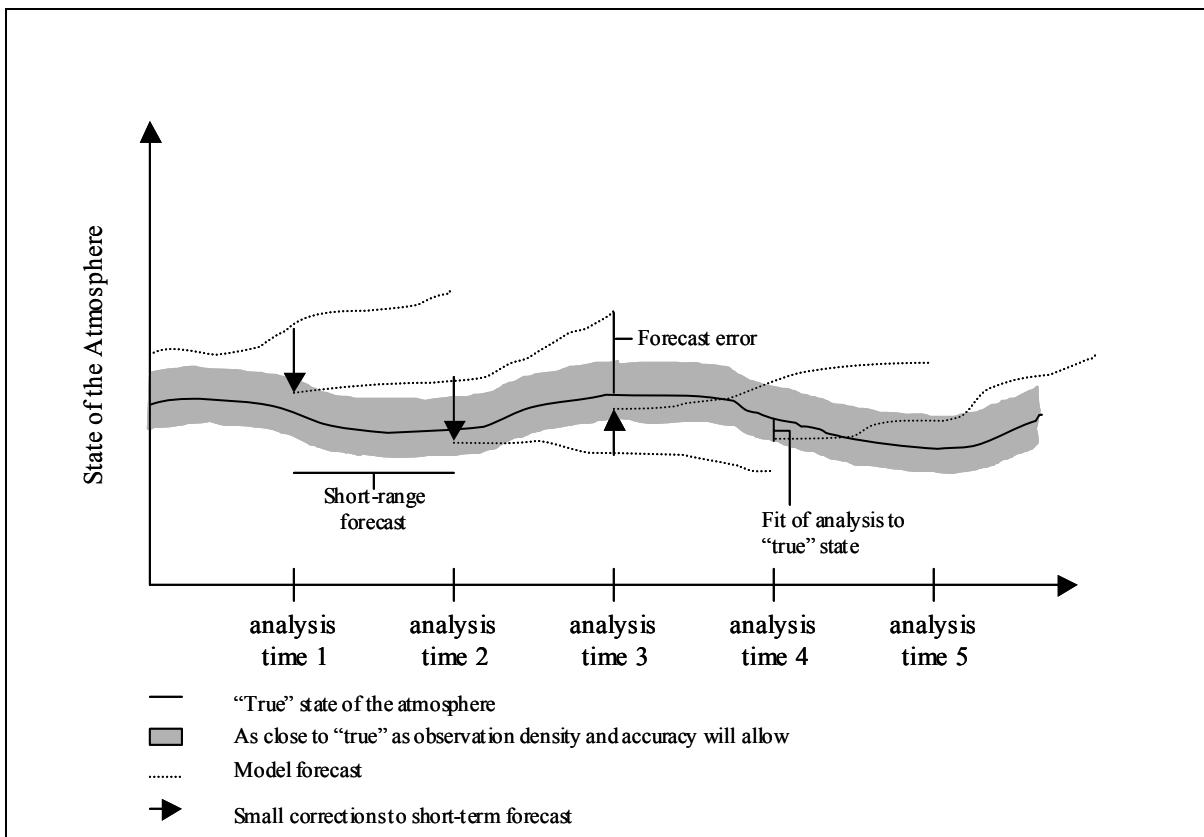
Figure 5. Data assimilation corrections. Observations are assimilated to correct each short-range forecast that serves as the basis for the next analysis (adapted from COMET, 2001).

2.4.2. The Data Assimilation Process. The data assimilation process results in a series of small corrections to the model forecast as shown in Figure 5. Ideally, the model forecast would be corrected to the "true" state of the atmosphere within the physical limits of

17

the model at each time. Since the "true" state of the atmosphere is never actually known, each analysis can only be as accurate as the observations and their physical limitations allow (COMET, 2001).

The data assimilation process involves ingesting data, decoding the observations, performing quality control to eliminate bad data, comparing the data to the model's first guess, and interpolating the data onto the grid as defined by TERRAIN. The process blends information from the short-range forecast with information from the new observations. This blending is the primary challenge in data assimilation. Observations must be transferred from numerous locations and times to the model grid, while at the same time preserving the physical, dynamical, and numerical consistency in the short-range forecast. The steps in the data assimilation process are as follows:

- Observations are ingested from outside the cycle.

- The observations are then checked for gross errors.

- Next, they're merged with a previous short-range forecast. This forecast is referred to as the "first guess" because it serves as the background for the next analysis. This merge is accomplished by computing observation increments, which are the differences between the observations and the first guess at the observation location.

- Quality control of the observation increments is then performed.

- An objective analysis procedure is used to interpolate the observation increments back to a model grid to produce a grid of model corrections.

- The corrections are then added back to the first guess to produce the new analysis.

- The new analysis provides initial conditions for the next operational forecast and a new short-term forecast to be used as a first guess to blend with the next set of observations.

This cycle repeats itself and is the basis of the data assimilation process. For a more in-depth look at data assimilation, the reader is referred to the web based COMET training module (COMET, 2001).

2.4.3. 3DVAR within the MM5 Modeling System. As previously mentioned, 3DVAR poses some changes in the way the data are preprocessed within the modeling system. Figure 6 is a schematic representing the flow of data from the MM5 preprocessing modules through the post-processor including 3DVAR. In Figure 6, all of the ovals represent either input or
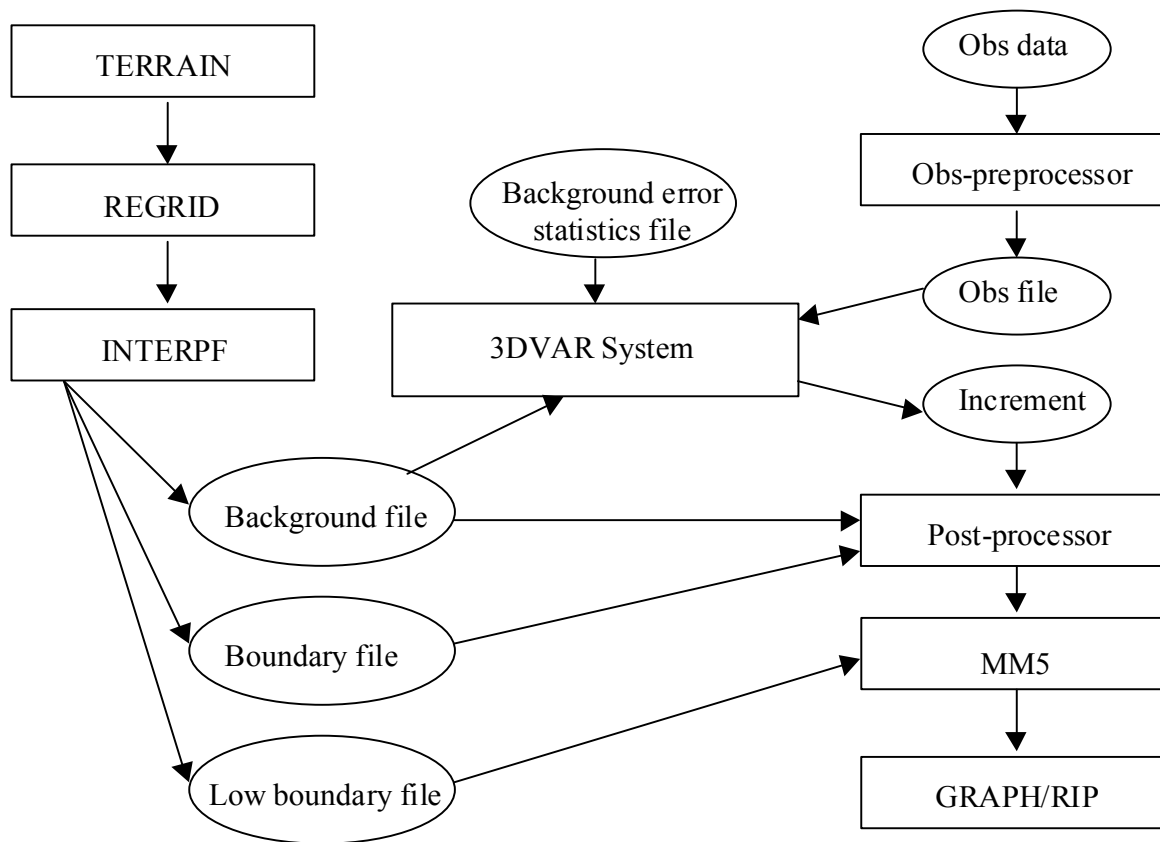
Figure 6. MM5 3DVAR system. Schematic diagram representing the flow of data through the MM5 modeling system using 3DVAR (adapted from NCAR, 2001).

output files used within the system. This configuration is one in which the output from

REGRID is input directly into INTERPF vice going into an assimilation program.

Additionally, the background file is input into 3DVAR where it is blended with the

observation file. The background file also goes into the post-processor, which is used to

update the boundary file. The increment file becomes the new background, or initial

condition, file used by MM5. All other processes are the same as described previously for

the MM5.

    2.4.4. 3DVAR versus MVOI. 3DVAR offers many advantages over MDAS/MVOI

as currently run in production by AFWA. The MVOI data assimilation scheme is relatively

inexpensive computationally and has been in use on and off at AFWA for the past 11 years.

As such, it is becoming antiquated as newer and better assimilation schemes are developed.

    The primary disadvantage of the MVOI system at AFWA is that the correlation

functions were designed for analysis on a coarse grid. Therefore, the MVOI cannot be used

to perform the analysis on fine scale grids (less than 45 km) directly. In order to perform

analysis on fine scale grids, the first guess from the MM5 non-hydrostatic sigma layers must

be interpolated to a hydrostatic sigma system. The analysis is performed on the hydrostatic

sigma surfaces, then it is interpolated back to the original non-hydrostatic coordinates. Each

step results in a small error that may grow throughout the forecast period (Ritz, 2001).

    Another disadvantage of the MVOI system is that it only provides height-based

analyses. This means that the temperatures are calculated based on height observations

rather than temperature observations and the wind field is geostrophically adjusted to match

the height field. This type of analysis can lead to significant errors in the upper air wind and

temperature fields. Also, with the MVOI, AFWA can only use the retrievals, not radiances,

from satellite soundings. This is not the best use of satellite data and the error associated with retrievals can be relatively large (Ritz, 2001).

Unlike 3DVAR, the MVOI system cannot directly assimilate satellite radiances, which have smaller errors associated with them than satellite retrievals. Satellite data are an invaluable resource and provide critical coverage, especially in the oceanic regions. Also, since the MVOI in use by AFWA is somewhat outdated, it is unable to make use of fields such as precipitable water. MVOI is also unable to take advantage of continuous data assimilation as described in Section 2.4.2. The 3DVAR system will be able to continuously assimilate data once it becomes operational at AFWA (Ritz, 2001).

The 3DVAR system is computationally expensive. Each 3DVAR analysis can take up to five times as much computing power as a one-day forecast for the same area (COMET, 2001). Although computational intense, 3DVAR provides a finer analysis of the meteorological conditions. It is able to use non-traditional observation types such as satellite radiances and precipitable water. AFWA is incorporating some of the advantages of MVOI in 3DVAR, such as the gross checking of observations (Ritz, 2001).

III. Methodology

3.1. Overview

The main focus of this research is to quantify the accuracy of the MM5 model forecast output using the 3DVAR data assimilation scheme with denied or limited input observations to the system. This investigation uses near real-time data due to the method of data collection, which will be described later in this chapter. Data collection commenced on 18 October, 2001 and ended on 14 December, 2001. The current production MM5 model run by AFWA and a developmental version of 3DVAR determined the design of this experiment. The MM5 model, as run for this experiment, was configured as similarly as possible to AFWA's operational MM5.

This chapter outlines the methodology and processes used in the data collection and the verification process of this study. First, there is a discussion of the computing requirements and resources used. The next section provides a description of the data collection process. Section 3.4 details the model configuration as used in this research, and section 3.5 is a description of the verification process.

3.2. Computing Resources.

The computing platform used in the production MM5 at AFWA is the IBM SPII supercomputer. Great effort was made throughout this research to duplicate AFWA's production MM5 as much as possible, including the computing platform. Therefore, all 3DVAR processes and MM5 model simulations were performed on an IBM SP P3 supercomputer located at the Aeronautical Systems Center Major Shared Resource Center

(MSRC) at Wright-Patterson AFB. All preprocessing (through INTERPF) and post-processing (after MM5) was accomplished on Sun Ultra 10/80 computers located at AFIT except as noted in section 3.3.

In an effort to maximize the use of the initial 1000 hours of processor time allotted for this project, a comparison was made and the most cost-effective number of processors to use for the MM5 runs was determined to be eight. Table 1 shows the comparison of processor expense. Ultimately, the initial 1000 hours requested to complete this project was a gross underestimate and approximately 2600 processor hours were used in all.

Table 1. Comparison of processor expense.

# of Processors	Time per Step (s)
8	2.79
16	1.51
24	1.03

3.3. Research Data

3.3.1. Input Data. The specific input data required were principally dictated by the goal of this thesis. 3DVAR and MM5 input requirements were also taken into consideration. AFWA provided all necessary input data files. Rather than providing the entire hemispheric AVN model forecasts, they provided the REGRID output starting at the MM5 model initialization time twice daily (00 and 12Z) for a 48-hour period. These REGRID output files provided all initial and boundary conditions for the full 48-hour period. AFWA also provided the observation files for each initialization time required by 3DVAR for the assimilation process as described in Chapter 2. Four observation files were available for each initialization time: surface observations, upper air observations, aircraft observations, and satellite observations.

3.3.2. Data Assimilation. Collecting data presented some interesting concerns due to the nature of this research. When considering the effects of denied or limited input data on model accuracy, a significant problem arises because AFWA initializes the MM5 using a short-range AVN model forecast. It was necessary to devise a scheme to work around the issue of the first guess forecast using all available observations in its assimilation since it is beyond the scope of this research to also deny the input observations to the AVN. The phrase "all available observations" refers to the observations made available to AFWA at the time of initialization in the context of this research. This project was run in near real-time and was designed to duplicate AFWA's production MM5 to the greatest extent possible, therefore the same input data were used that goes into the production MM5.

The scheme devised to work around the issue of the observations being allowed into the AVN model takes full advantage of the 48-hour REGRID output files provided by AFWA. The premise behind the entire data collection process was that each model run, whether it contained all available observations or had denied or limited input data, was for a 24-hour period. The first 24-hour period of data in these files was used to provide the initial and boundary conditions for model runs in which all available observations were to be assimilated. The last 24-hour period of the REGRID file from the previous day supplied the initial and boundary conditions for runs in which any or all of the observations were to be denied. Figure 7 is a timeline detailing the process of data input to the system.

The method of data assimilation used in no way eliminates the observations from being allowed into the earlier AVN model runs. It does, however, provide a means of allowing the observations to grow older and have less impact on the assimilation process and ultimately on the output. Conversely, it allows the first guess field used by 3DVAR to be as
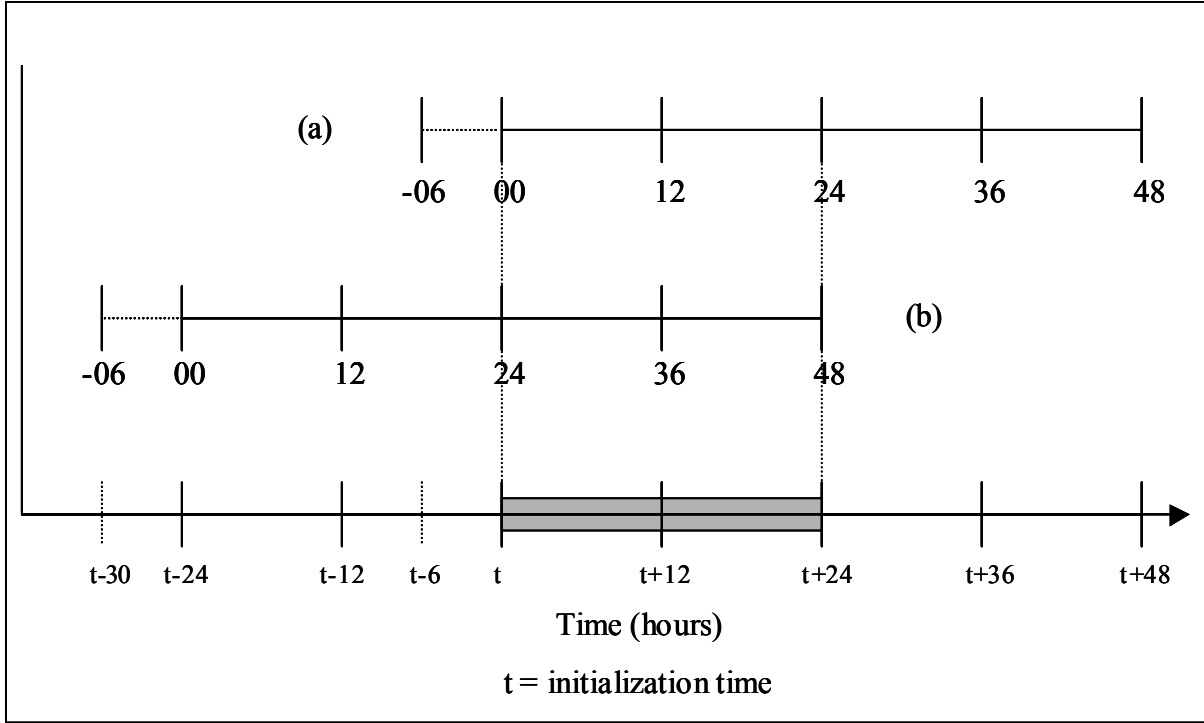
Figure 7. Use of AVN forecast for model initialization. Timeline depicting how AVN model forecasts are used to provide initial and boundary conditions. Line (a) represents the REGRID data file from which the initial and boundary conditions come for the case with all observations and line (b) represents the REGRID file used for the data denied cases. The shaded portion of the timeline at the bottom is the forecast period. The –06 hour mark on lines (a) and (b) indicates the AVN initialization time used to create the REGRID files.

current as possible for the case where all available observations are used. It is important to note that AFWA uses various times of AVN and NOGAPS model runs to initialize the MM5. To ensure consistency throughout this project, only 6-hour old AVN forecasts were used as a first guess for the cases where all available observations were used. This means that for each data denied case, the first guess field was at least 30 hours old. This architecture allowed the older observations to have less impact on the overall outcome of the model forecast.

The use of different first guess fields and boundary values from the various REGRID files for model initialization presented a significant concern in that different initial and boundary conditions were used in the control runs than in the data denial experiments. The errors associated with these different conditions must be accounted for when comparing the

data denied cases with the control group. An additional set of control runs was performed for this purpose as described in the following section.

3.3.3. Test Configuration. Several different 3DVAR configurations were used to acquire the various required MM5 model output files. Each model run in which all available observations were assimilated served as a control group, or baseline, to which all other cases for the same initialization time were compared. Five runs were performed for each initialization time: all available observations used with the 6-hour old AVN forecast (C1), all available observations used with the 30-hour old AVN forecast (C2), total observation denial (TOT), upper air observations denial (UPP), and surface observations denial (SFC). Table 2 provides a list of the different experimental runs and describes the configuration of the first guess field and the input observations used for the various cases. The naming convention given in the table is used throughout the analysis portion of this study.

Table 2. List of experiments. The main control run, designated C1, uses all available observations and the 6-hour old AVN forecast for the first guess field while C2 is the secondary control run used to estimate errors resulting from the use of different initial and boundary conditions.

Experiment name	First guess	Denied observations	# of runs per exp
C1	6-hour old AVN	None	00Z - 37
			12Z - 26
C2	30-hour old AVN	None	00Z - 9
			12Z - 10
TOT	30-hour old AVN	All	00Z - 37
			12Z - 26
UPP	30-hour old AVN	RAOB only	00Z - 37
			12Z - 26
SFC	30-hour old AVN	Surface only	00Z - 37
			12Z - 26

The observation files containing the satellite and aircraft data were not individually denied. The satellite data were not denied because the only information currently available for the European theater is cloud-derived winds (Ritz, 2001).

The second control run (C2) was used to estimate the errors associated with the use of different initial and boundary conditions between the main control run (C1) and the data

26

denied test cases (TOT, UPP, and SFC). In a real-world data denial situation, the observations would not be available for assimilation by either the MM5 or AVN models. This being the case, initial comparisons of C1 to the various test cases were made in an effort to quantify the overall effects of data denial and limitation on the MM5 model output. A second set of analyses was made by comparing C2 to the various data denied test cases, which indicates the errors of only the data denial since C2 and all the data denied cases use the same initial and boundary conditions. The new analyses created were then subtracted from the initial comparisons made using the C1 control group to estimate the errors created by changes in the first guess field from the primary control run. Without the secondary control group and subsequent analyses, two variables (observation inputs and first guess field) would have been changed and the source of errors could not be definitively linked to the data denial.

As shown in Table 2, a different sample size was used for the C2 control run. The primary reasons for the different sample size were the computational expense of performing this experiment and the fact that it is only used to estimate the differences created by the use of different initial and boundary conditions between the C1 control group and data denied test cases. It is important to note that during all analyses with C2, only the data created from the same initialization dates and times were used to make comparisons.

3.4. Model

The focus of this research was to quantify the effects of denied/limited input data on the MM5 model forecast accuracy using 3DVAR for data assimilation, thus dictating the

27

overall design and setup of the experiment. This section outlines the various model

configurations used throughout the course of the current research project.

  3.4.1. Model Configuration. The MM5 is a versatile modeling system allowing

many different alternatives to modify the domain, parameterization schemes, and numerous

other variables necessary for the model simulations. Selection of the configuration for each

parameter depends on the particular requirements of the user. All parameters were set to

those used by the AFWA production MM5 except the time step and forecast length as shown

in Table 3.

Table 3. MM5 model configuration.

Parameter	AFWA	Current
Moisture	Reisner 1	Reisner 1
Cumulus	Grell	Grell
PBL	MRF	MRF
Radiation	Cloud	Cloud
Landuse Data	USGS	USGS
Time Step	99 s	120/90 s
Forecast Length	72 hr	24 hr

  The Mulitprocessor Instruction (MPI) version of the MM5 was used to perform all

model runs after the data were assimilated in 3DVAR. The MPI version allows user input of

the number of processors between which the MM5 will distribute the domain. The domain is

split between the north-south and east-west directions. For this research, two processors

were used in the north-south direction and four processors in the east-west direction. The

MM5 output data were transferred from the MSRC back to the AFIT Meteorology Lab for

post-processing on the Sun Ultra 10 after the model was run for the four different input

configurations.

  The initial configuration was set up with a 120-second time step for this project in an

effort to conserve the limited computing resources available. However, this time step was

too large and violated the Courant-Freidrichs-Lewy (CFL) stability criterion for some experiments. Equation 2 describes the CFL criterion where Δt is the time step, Δx horizontal grid spacing, and c is the wave speed.

$$\Delta t \le \frac{\Delta x}{c} \tag{2}$$

The CFL criterion mandates that there is a minimum time step that can be used for stability in an explicit time differencing schemes, and it is a function of both the horizontal grid space and the wave speed of the solution (Grell et al., 1994). Using a fixed value for the horizontal grid space, the time step becomes a function of the wave speed; the faster the wave speed, the smaller time step necessary for stability. When this stability criterion is violated, the numerical scheme becomes unstable and the numerical output becomes unusable.

Several model runs during the period of 31 October and 10 November 2001 violated CFL and had to be rerun using the smaller time step. The new time step of 90 seconds was chosen to ensure CFL was not violated again in the future since the speed of the fastest wave cannot be determined ahead of time. To ensure consistency with the previous data, a single model run was made using both the 120-second time step and the 90-second time step in order to make a comparison of the output. The results of the surface data comparison are found in Table 4. The differences in the output data of the two runs were determined to be insignificant. Therefore, the time step was changed partway through the data collection process, and the runs that violated the CFL criterion were reaccomplished.

Table 4. Comparison of surface data using 120-second versus 90-second time step.

Temperature	Average RMSE			BIAS		
Forecast Hour	120s	90s	Difference	120s	90s	Difference
00	1.8244	1.8244	0.00	-0.073	-0.0731	0.00
03	2.0134	2.0111	0.0022	-0.147	-0.1399	-0.0071
06	2.1452	2.1674	-0.0222	-0.278	-0.2734	-0.0047
09	2.8375	2.8262	0.0114	-1.821	-1.807	-0.0137
12	3.0706	3.0525	0.0181	-1.978	-1.9549	-0.0228
15	2.5446	2.5103	0.0343	-1.585	-1.5582	-0.0267
18	2.3187	2.3113	0.0074	-0.356	-0.3205	-0.0355
21	2.232	2.257	-0.025	0.0854	0.0967	-0.0113
24	2.2754	2.3065	-0.0311	0.4679	0.4821	-0.0141

Pressure	Average RMSE			BIAS		
Forecast Hour	120s	90s	Difference	120s	90s	Difference
00	3.7118	3.7118	0.00	2.4097	2.4097	0.00
03	4.4445	4.4505	-0.006	1.9894	2.0167	-0.0274
06	4.2367	4.2416	-0.0049	1.6375	1.6487	-0.0111
09	3.9771	3.9785	-0.0015	1.2977	1.3104	-0.0127
12	4.2709	4.2802	-0.0094	1.8426	1.8621	-0.0195
15	4.3829	4.3886	-0.0057	2.0422	2.0551	-0.0129
18	4.3472	4.3572	-0.01	1.781	1.7966	-0.0156
21	4.4778	4.4849	-0.0071	2.1365	2.1542	-0.0177
24	3.9628	3.9779	-0.0151	2.3319	2.3651	-0.0332

Wind Speed	Average RMSVE		
Forecast Hour	120s	90s	Difference
00	3.6174	3.6174	0.00
03	3.9685	3.9699	-0.0014
06	3.9307	3.937	-0.0063
09	3.4958	3.5059	-0.0102
12	3.3973	3.4083	-0.011
15	3.6945	3.6917	0.0029
18	3.7857	3.8037	-0.018
21	3.8277	3.8172	0.0105
24	4.1408	4.1117	0.0291

3.4.2. Model Domain. The European theater was chosen as the domain due to the fact it was in this theater that the 3DVAR was first in the testing phase at AFWA. Figure 8 illustrates the domain including the relative location of all surface and upper air observation sites used for verification purposes. It is important to note that approximately one third of the surface area for this domain is water. It is virtually impossible to verify the model output over the water because there are very few reporting stations available for surface

observations and no ground-based RAOB stations. Satellite soundings could not be

assimilated into the version of 3DVAR used for this study.
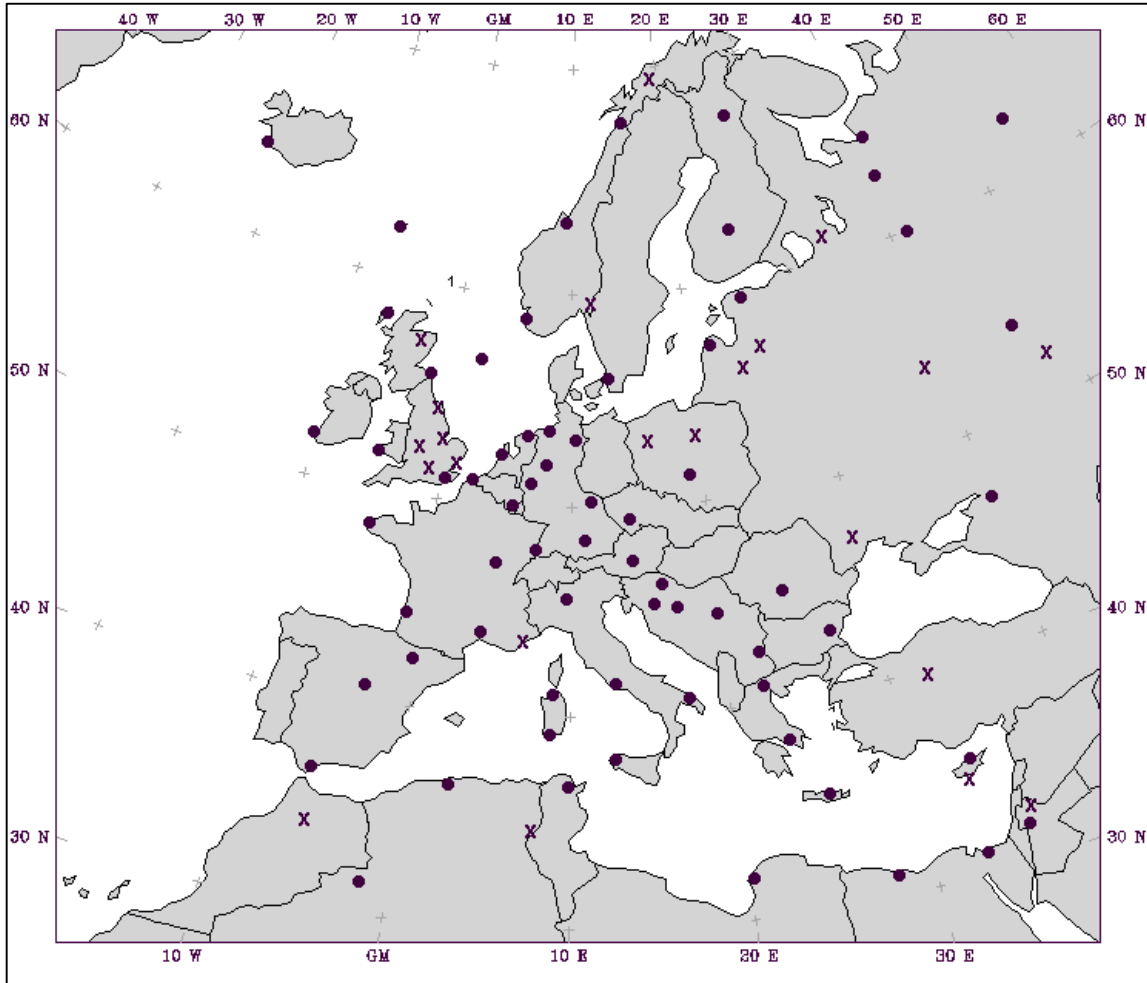
Figure 8. Model domain. This diagram shows the boundaries on the surface of the entire European window and the relative location of all verification stations. The stations indicated by (**x**) were those only providing surface observations while the ones indicated with (●) provided both surface and upper air observations for verification purposes.

3.5. Verification Procedure

     3.5.1. Overview. Each 24-hour forecast generated an output file totaling

approximately 220 megabytes of data. An extensive amount of post-processing was required

to convert the MM5 output into a more useable form. The program used to manipulate the

raw model output into useable data sets was RIP. The RIP program converts the model generated output parameters into a FORTRAN readable binary format. The RIP output was used to accomplish both the surface and vertical verifications throughout the course of this study.

Verification is the process of comparing the forecasted model output to surface observations or atmospheric sounding data. AFWA uses on the order of 2000 surface stations for verification of the European window. It is beyond the scope of this study to attempt a verification of that magnitude. Ninety surface stations were chosen for the purpose of verification in this research. Several criteria were used to select the surface stations for verification. First, stations were required to be on the list of stations verified by AFWA. The selected stations were also limited to those reporting both surface and upper air observations. Finally, the stations were chosen to provide a relatively uniform distribution of locations throughout the domain. Only 68 of the 90 stations reported any RAOB data during the period of this study, thus limiting the number of upper air stations to 68 while still allowing 90 surface verification stations. See Figure 8 for a graphical representation of the relative locations of all verification stations. Appendix A contains a list of all verification stations used throughout this study.

A grid-to-station verification method was used to compare the model output to the station observations for both surface and upper air observations. An inverse-linear interpolation and weighting technique was used to compute the model value at the exact station location on the model grid, shown in Figure 9. The closest grid point to the station contributed the greatest influence while the other surrounding grid points had less impact.

The same horizontal interpolation and weighting procedure was used for both the surface and upper air verifications.
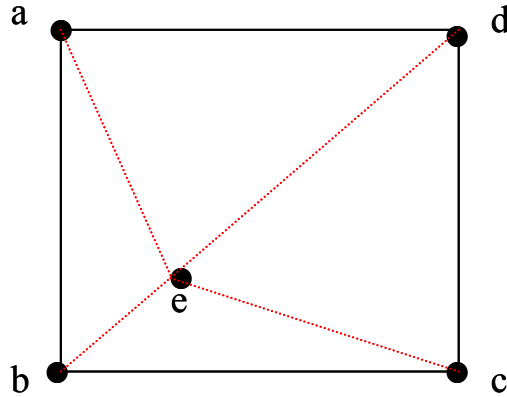

Figure 9. Grid-to-station interpolation. Inverse-linear interpolation was used to compute the model value at the verification station (e). This method allows the most weight to be given to the model grid point (b) closest to the station while the others still have some influence.

3.5.2. Vertical Verification Process. RIP was used to extract vertical parameters of geopotential height, pressure, temperature, water vapor mixing ratio, cloud ice mixing ratio, cloud water mixing ratio, rain mixing ratio, snow mixing ratio, east-west wind component, and north-south wind component. Each of the vertical parameters from the RIP output is interpolated horizontally for each of the 68 RAOB reporting stations on all 41 half-sigma levels. The RAOB data contained all mandatory and significant levels, which varied for every station and every observation time. These data were logarithmically interpolated from pressure surfaces to the model sigma levels vertically. Interpolating logarithmically in pressure coordinates proved to be more accurate than linear interpolation due to the nonlinear vertical decrease of pressure in the atmosphere.

Interpolation was conducted from RAOB pressure to sigma level for two reasons. The focus on low-level and surface interactions dictates a higher model resolution near the

surface. The higher resolution required near the surface is evident in Figure 2 where the sigma levels are closer together in the lower atmosphere. Conversely, there are fewer RAOB pressure levels of observations on average than sigma layers in the lower levels. Second, the number of RAOB levels reported fluctuates based on balloon reporting. This fluctuation leads to a varying number of levels being reported by the RAOB for each location and time.

The horizontally interpolated model data are then subtracted from the vertically interpolated RAOB data to arrive at a residual on the same pressure level. The residuals were calculated using equation 3.

$$residual = model\ output - observation \tag{3}$$

The convention used in equation 3 allows a positive residual to be interpreted as a model over-prediction. The residuals for each station are averaged for all 00Z runs and all 12Z runs individually. Due to diurnal variations, the separation of 00 and 12Z runs was required to eliminate statistical errors inherent in data dependence. The overlapping of forecast periods resulted in a violation of independence, especially for the small sample sizes of this study. The sample sizes for the respective initialization times are 37 for 00Z and 26 for 12Z.

Following Wilkes (1998), the statistical parameters of root mean square error (RMSE) and bias were computed in order to make the necessary comparisons between the model outputs. The residuals were averaged throughout the vertical profile to arrive at a mean residual, or average bias. Equation 4 was used to calculate the RMSE, where *residual* was calculated in equation 3 and *n* is the number of valid observations.

$$RMSE = \sqrt{\frac{\sum (residual)^2}{n}} \tag{4}$$

A paired *t*-test of the residuals was conducted on the model output to determine vertical levels that were statistically different from the observations at a 95% significance level. An assumption of normality had to be made in order to use the paired *t*-test. P-values were calculated in order to determine if there was a statistically significant difference between the compared values. The difference ($\Delta = allobs - denied$) was used as the basis for all verifications. Therefore, the null hypothesis H_0: $\Delta = 0$ was used in every case. The 95% significance level corresponds to a type-I error (α) of 0.05. The null will be rejected in favor of the alternative, $\Delta \neq 0$, when the P-value is less than or equal to α. Rejecting the null indicates there is a statistical difference between the two data sets. When the P-value is greater than α, the null cannot be rejected (Devore, 2000).

3.5.3. Surface Verification Process. The RIP data processing program is used to extract the lowest sigma level parameters of temperature, pressure, east-west wind component, and north-south wind component. The lowest sigma level is located at constant 20 meters above ground level (AGL) throughout the domain. All surface temperature observations are taken at a standardized 2 m AGL. The surface temperature is calculated at the lowest half-sigma level, which is 20 m above the surface. Therefore the model does not produce a surface temperature directly. However, a ground temperature variable is computed. The ground temperature, also known as skin temperature, is not the same as the surface temperature as measured at 2 m AGL in the observation process. A simple random check against the ground temperature values proved that it was inappropriate to verify the model skin temperature against the observed surface temperature. The ground temperatures were considerably higher than the observations. The lowest sigma level values for

temperature were compared to the observations taken at 2 m AGL without any extrapolation or corrections in order to verify the surface temperature parameter.

The inverse-linear interpolation and weighting process previously described was utilized to arrive at model output values for the specific station locations on the model grid. The verification procedure was performed at 3-hour intervals starting with the initialization time and running through the 24-hour forecast. Residuals of the parameters were calculated following equation 3.

The surface parameters verified in this study were temperature, mean sea level pressure, and winds. RMSE and bias were both computed for all stations and used in the temperature and mean sea level pressure verifications. Root mean square vector error (RMSVE) was used in the verification of the surface winds. RMSVE was chosen because it takes both wind speed and direction into account, thus eliminating the need for cumbersome direction comparisons. RMSVE is calculated using equation 5,

$$RMSVE = \sqrt{\frac{\sum[(u_{mod} - u_{obs})^2 + (v_{mod} - v_{obs})^2]}{n}} , \qquad (5)$$

where u and v denote the east-west and north-south wind components respectively, the subscripts mod and obs are the model and observed parameters, and n is the number of valid surface observations for a specific forecast time. The paired t-test as described in Section 3.5.2 is used for the surface verification after the RMSE, RMSVE, and biases were computed.

The model pressure at the lowest half-sigma level required a correction factor of 2.16 mb in order to reduce it to the surface. The correction factor was determined by applying a hydrostatic approximation to the height difference between the lowest half-sigma level

(20 m) and the observation height (2 m). This value is constant because the height difference

(18 m) remains constant throughout the domain.

3.6. Terrain Analysis

Prior to verifying the surface parameters, it was necessary to conduct a terrain

analysis. The analysis is a necessary prerequisite because surface parameters are highly

influenced by the topography. The European theater contains many various types of

topography ranging from the low, relatively flat ocean surfaces to high peaks of the Alps and

other mountain ranges. Because the grid resolution for this study was 45 km, this warranted

a critical assessment of the accuracy of the terrain grid used in the simulations.
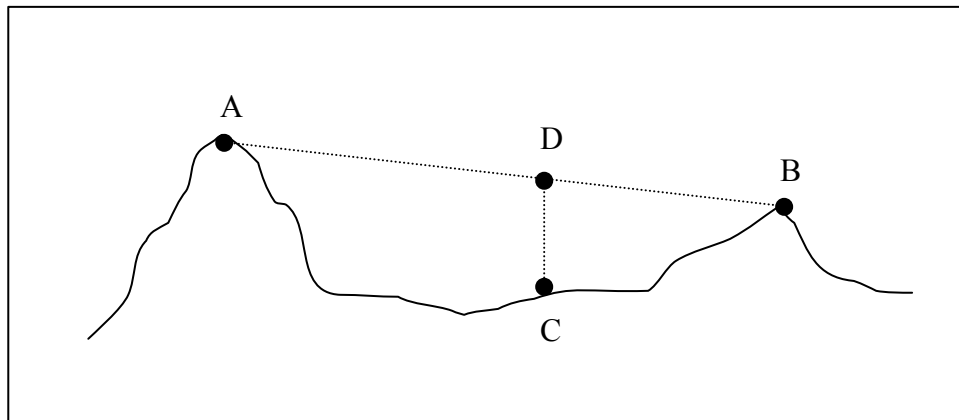
Figure 10. Calculation of terrain residual. A two-dimensional representation of the method of computing
terrain residual. Points A and B represent model grid points. Point C is the verification station. Point D
represents the elevation as computed by the model for the given station. The difference in elevation between
points C and D is the terrain residual.

The same interpolation and weighting functions that were used to compute the

horizontal verification results were applied to the model terrain grid. Figure 10 shows a

crude two-dimensional example using only two grid points of how the model terrain values are computed for each station.

The residuals were consistently positive, indicating an overestimate of actual station elevations. It was necessary to exclude certain stations exceeding a residual criterion. These findings are consistent with tests conducted by Manning and Davis (1997). They also found the average model terrain to be too high from the smoothing of the grid resolution.

The distribution of lowest sigma level temperatures was analyzed to determine a residual criterion in order to exclude certain surface verification stations. The distribution of the lowest sigma level temperatures should ideally compare to a terrain residual distribution. It was found that approximately 99% of these temperature residuals were less than 3.5 K. This was defined as the acceptable range of error based on a dry adiabatic lapse rate. Assuming the dry adiabatic lapse rate of 10 K/1000 m, 3.5 K corresponds to a terrain residual criterion of 350 m. Four stations with an absolute value of terrain residual greater than 350 m were excluded from the surface verification process. The excluded stations can be seen in Appendix A.

IV. Analysis and Results

4.1. Overview

This chapter presents the results of the statistical analyses performed in this study. There were three different data denied cases for each model run, thus there are three sets of charts for each initialization time and parameter analyzed. The chapter starts by discussing the upper air verification results, followed by a discussion of the results of the surface verification. A final comparison is then made of the results of the analyses using the primary control group (C1) and those of the analyses using the secondary control group (C2) accounting for differences in the initial and boundary conditions resulting from using different AVN forecasts for the primary control group and data denied test cases.

All observations used for verification were provided by the Air Force Combat Climatology Center (AFCCC) and are assumed to be error-free. However, a rough quality control routine was performed on all observations to remove data with unrealistic values. Unrealistic values are defined as temperatures greater than 50°C, pressures above 1040 mb, and winds greater than 100 knots. In addition, the accuracy of the observations is not known and is assumed to be within the limits of the equipment used to measure each parameter. The specific observing equipment used in the European theater is assumed to be of comparable quality and accuracy as that used by the National Weather Service and the U.S. Air Force. Table 5 lists observation equipment accuracy as advertised in information published for the various instruments. The surface sensor accuracies are based on the standard equipment found in Air Force weather stations worldwide, while the upper air data are based on the accuracy of the Vaisala RS80 family of radiosondes used in Europe (BADC, 2002).

Table 5. Observation equipment accuracy (from Vaisala (2002) and Department of the Air Force (1995, 1997, and 1998)).

Surface Observations		
	Parameter	Sensor Accuracy
	Temperature	± 1.0 K
	Pressure	± 0.5 mb
	Wind Speed	± 1.0 m/s
	Wind Direction	± 5.0°
Upper Air Observations		
	Parameter	Sensor Accuracy
	Temperature	± 0.2 K
	Pressure	± 0.5 mb

All differences were computed using the convention that the difference (Δ) equals the control group (C1 or C2) minus the data denied test case (TOT, UPP, or SFC) to ensure consistency. This method of differencing allows positive differences to indicate the control group produced a higher value at that point for the particular parameter being analyzed. Likewise, a negative difference shows that the test case had a greater value. Refer to Table 2 for a description of the configuration of the various experimental runs used throughout this study.

4.2. Upper Air Verification Results

This section presents the findings of the upper air verification process. The upper air verification statistics were calculated using the 68 RAOB stations listed in Appendix A. The upper air parameters analyzed in this study were geopotential heights and temperature.

Weather balloons are generally launched every 12 hours, thus dictating the verification times used. Each 24-hour forecast covered three RAOB reporting times. These times correspond to the model initialization time as well as the 12 and 24-hour forecasts.

Very few balloons made it above the 100 mb level, thus providing little data to be analyzed at the two highest sigma levels. Therefore, no verification was performed at levels above 100 mb.

Paired *t*-tests, as described in Section 3.5.2, are used at each sigma level where the model output data are compared to the interpolated RAOB data. The significance tests used determine the statistical relevance of the differences in RMSE and bias separately. All upper air significance tests were performed using the null hypothesis H_0: $\Delta = 0$ and a 95% confidence interval. Use of the paired *t*-test defined the differences as statistically significant at nearly every level.

The black bars in all of the following charts show the results of the paired t-test. The presence of a black bar indicates that there are statistically significant differences between the compared model outputs at those points. The absence of a bar suggests that the differences between the parameters were not large enough to be statistically significant. A conclusion cannot be drawn, however, that the two model outputs are the same at that point. The only conclusion that can be made is that there is not enough evidence to prove they are different. When using a paired t-test with a 95% confidence interval, there is an inherent 5% chance that the test is wrong.

All RMSE charts presented in this section show only a single trace detailing the difference between the two model runs being compared. All bias charts have both profiles plotted. The difference in plotting between the RMSE and bias is primarily a function of the method of computing the errors. The average RMSE differences only indicate how close the model forecast outputs are to each other with respect to the observations. The RMSE differences are computed from the average RMSE values of the respective model output, and

are always a positive number. The bias is the value showing the tendency of the model.

Biases can be positive or negative depending on whether the model over or under forecasts

the parameter. Differencing the biases can act as a filter and remove valuable information

because these values can be either positive or negative, which is important in analyzing the

biases.

    4.2.1. Temperature Verification. Figures 11 through 13 show the temperature RMSE

differences between C1, which used all available observations, and the data denied test cases

(TOT, UPP, and SFC respectively) for both the 00 and 12Z initialization times on sigma

levels. Recall that negative values indicate areas where the data denied case generated

greater errors than the case using all available observations. The opposite is true for positive

differences.

    In all cases, the lowest sigma levels show differences comparable to those of the

surface analysis, shown in section 4.3.1. Figure 11 is a diagram depicting the RMSE

differences for the TOT test case. As one would intuitively expect, the test case showed a

greater error than the case where all available observations were used. It is important to note

that the horizontal scales differ between the model times. The tendency was for the RMSE

differences to decrease with time, while the actual RMSE values increase throughout the

forecast period. This result indicates that the observations used at the model initialization

time are important, but as the model gets further from the initialization time, they have less

impact on the forecast. This trend is not indicative of the model forecast getting better. It

does, however, show that the errors converge as the model gets further away from its
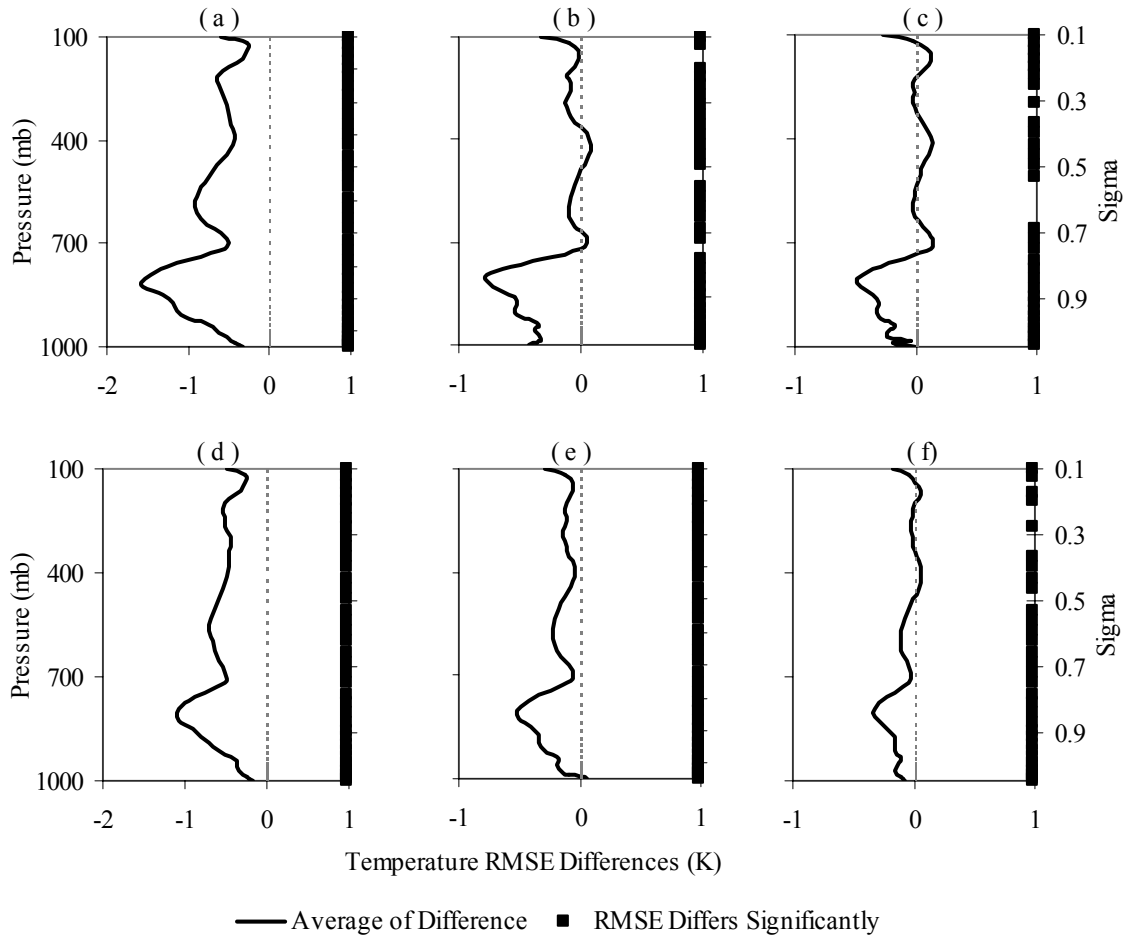
initialization time.

Figure 11. Average differences (C1-TOT) in upper air temperature RMSE. Charts (a), (b), and (c) represent the 0-hr, 12-hr, and 24-hr forecasts respectively for the 00Z initialization time. Charts (d), (e), and (f) represent the 0-hr, 12-hr, and 24-hr forecasts respectively for the 12Z initialization time. The black bar denotes sigma levels with a statistically significant difference at the 95% significance level.

Figure 12 shows the temperature RMSE difference for the UPP test case. An important feature to note is the contrast between these charts and those in the previous figure. The RAOB data denied test case initialized only slightly better than the TOT case, suggesting that the aircraft, satellite, and surface observations assimilated have minimal influence on the model initialization. Also of importance is that the RMSE differences converge as the model gets further from the initialization time, indicating that the boundary conditions are more important than the observations later in the forecast period.
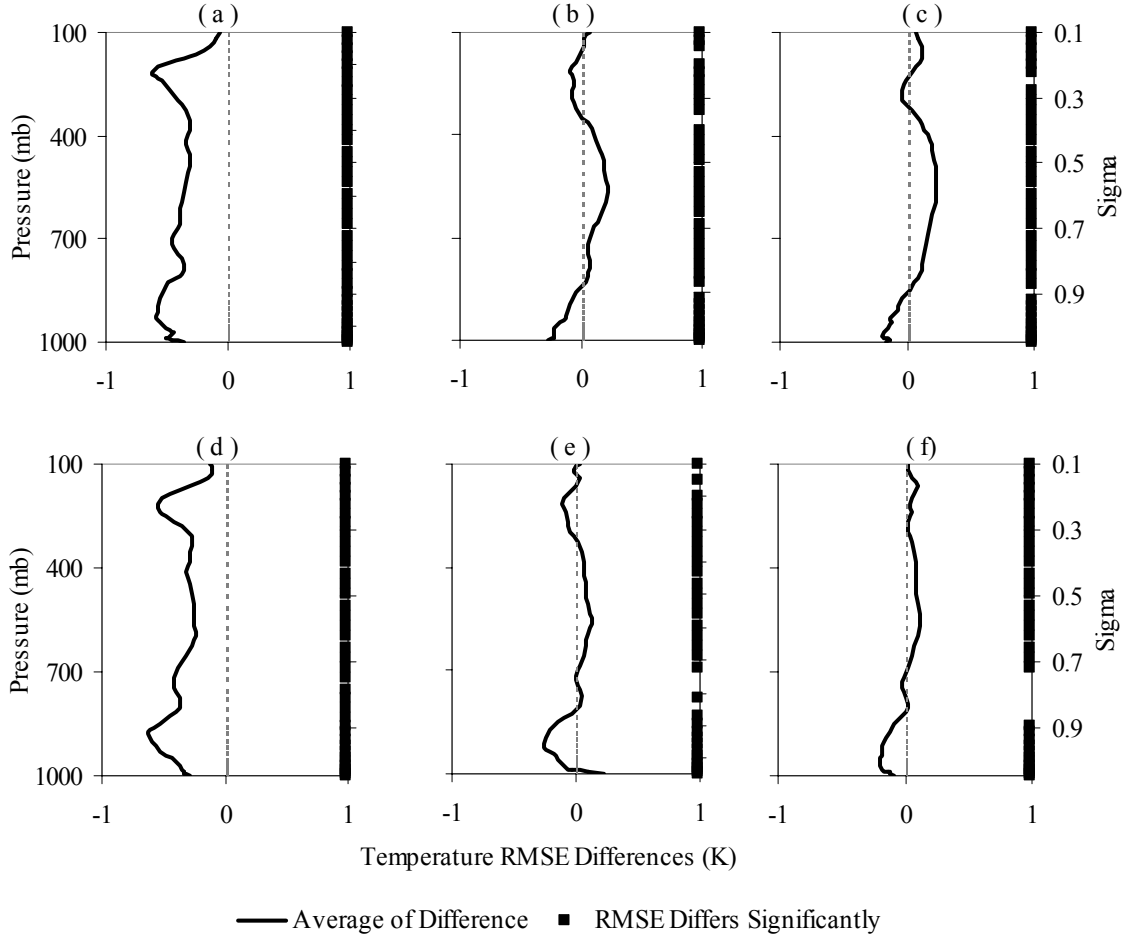
Figure 12. Average differences (C1-UPP) in upper air temperature RMSE. Same as Figure 11.

Figure 13 shows the vertical profile of the temperature RMSE difference for the SFC test case. The difference in RMSE at the initialization time was markedly better than either of the other test cases in all instances. This difference can be attributed to the fact that the lowest sigma level in the model is at 20 m AGL and the surface temperatures are measured at 2 m AGL. Although the surface observations are assimilated into the lower boundary conditions of the model, their influence is less important in the upper levels of the atmosphere.
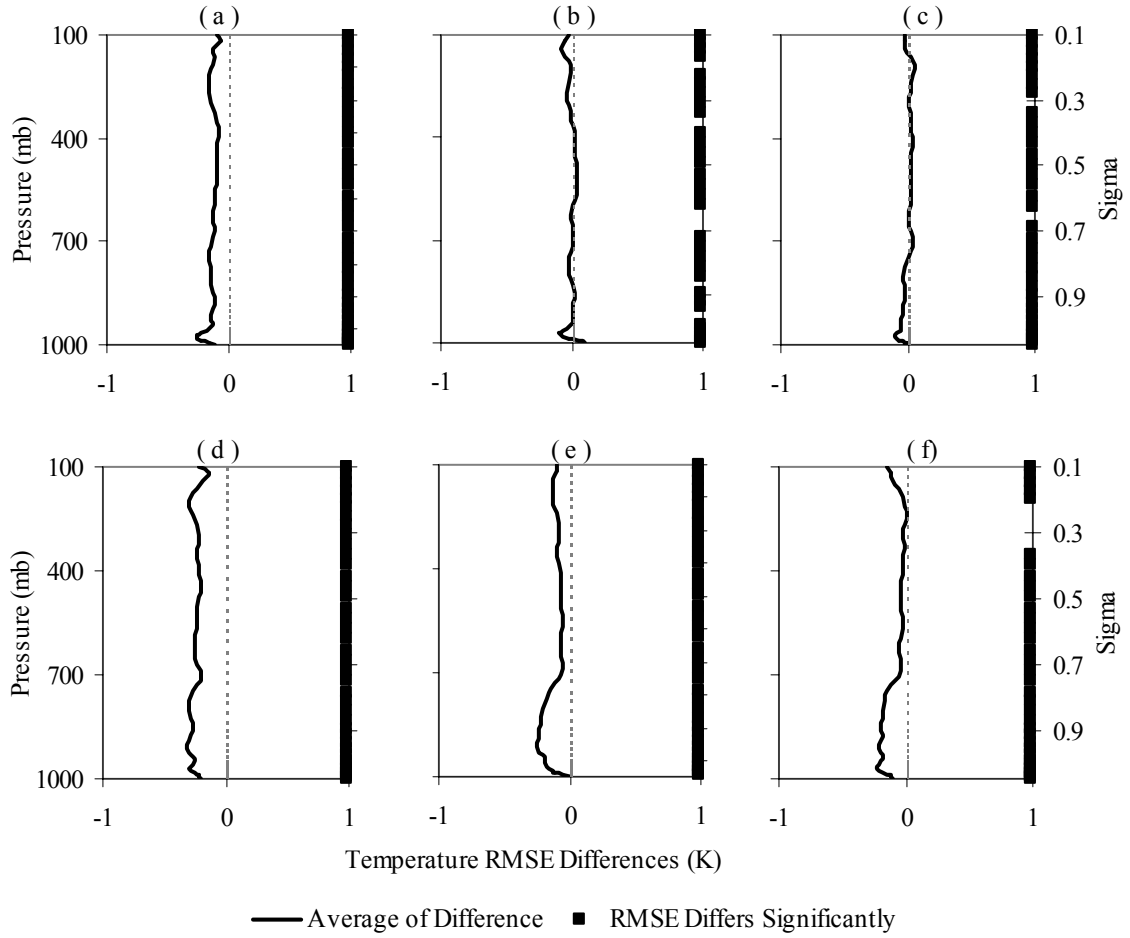
Figure 13. Average differences (C1-SFC) in upper air temperature RMSE. Same as Figure 11.

Figures 14 through 16 display the average temperature bias profiles in much the same manner as the RMSE differences were presented. The main difference is that both profiles are shown together rather than simply showing the difference for the reasons previously mentioned. In all cases, there is a marked difference between the 00Z and 12Z initialization times near the surface. This difference is a direct result of diurnal variation and is very pronounced in all temperature bias analyses of this study. This error is described in more detail in section 4.3.1.

The charts in Figures 14 and 15 show a comparison of the average temperature bias of the C1 control case versus that of the total and upper air data denied test cases respectively. These diagrams present similar analyses and will be discussed together. Near the surface, the biases are very similar, which is consistent with the findings of the surface verification shown later. However, ascending in the atmosphere, this similarity becomes less pronounced, indicating the importance of the observations in the data assimilation process.
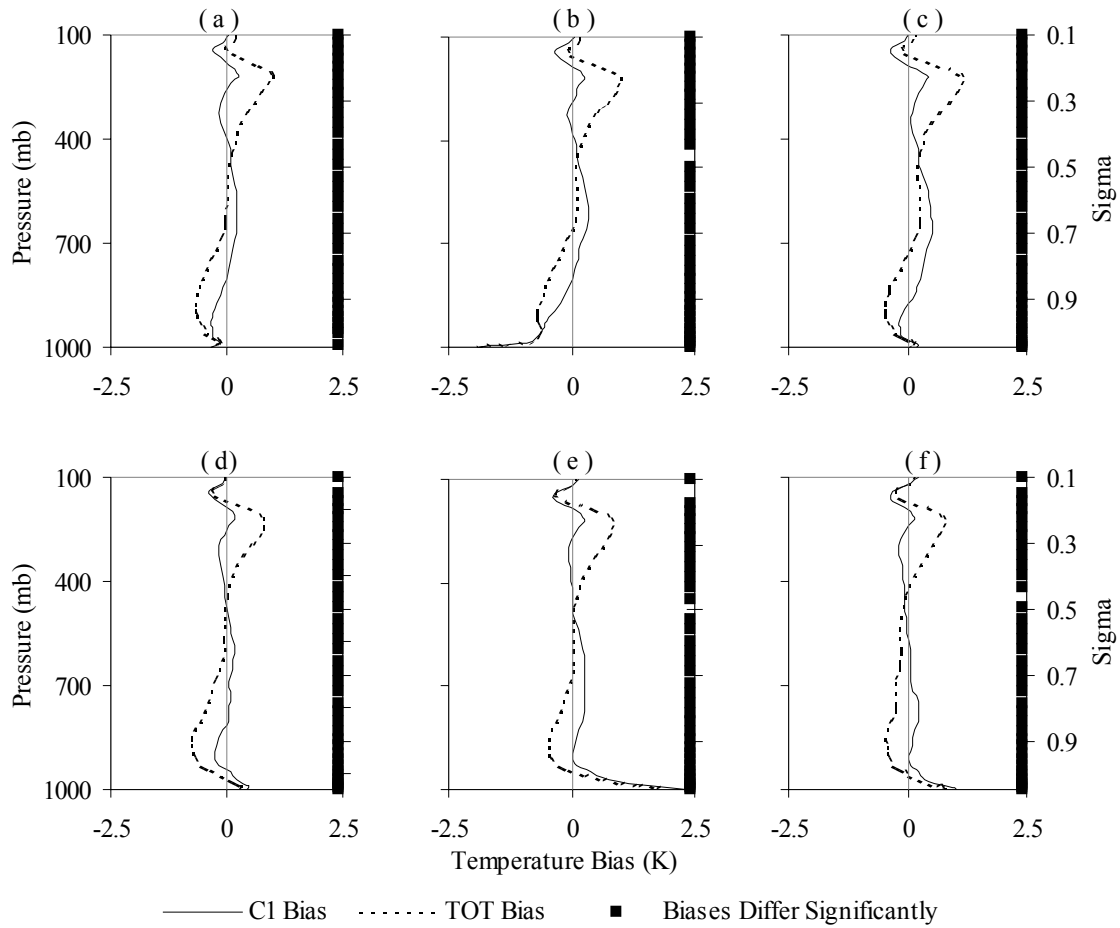

Figure 14. Average upper air temperature bias comparison (C1 vs. TOT). Charts (a), (b), and (c) represent the 0-hr, 12-hr, and 24-hr forecasts respectively for the 00Z initialization time. Charts (d), (e), and (f) represent the 0-hr, 12-hr, and 24-hr forecasts respectively for the 12Z initialization time. The black bar denotes sigma levels with a statistically significant difference at the 95% significance level.

In addition, the 00Z analyses shows a cold bias at the lower levels while there is a warm bias higher in the atmosphere.
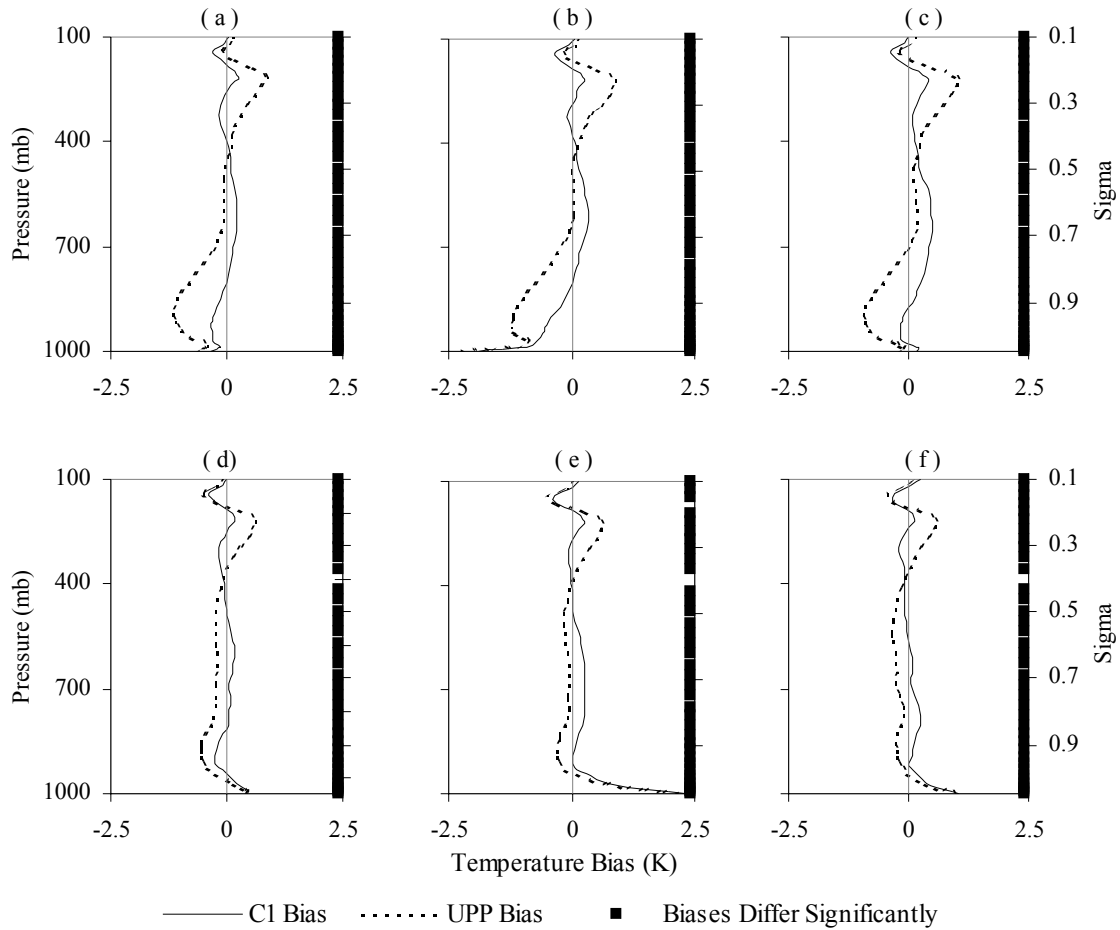

Figure 15. Average upper air temperature bias comparison (C1 vs. UPP). Same as Figure 14.

The analysis presented in Figure 16 shows the comparison between the C1 control group and the surface data denied test case. Unlike the total and upper air data denied test cases, denying surface observations has relatively little influence on the vertical temperature profile. This trait is evidenced by the SFC bias being considerably closer to that of C1 than for either the TOT or UPP test case indicating the importance of the upper air observations in

the data assimilation process. This finding is consistent with the analysis of the RMSE
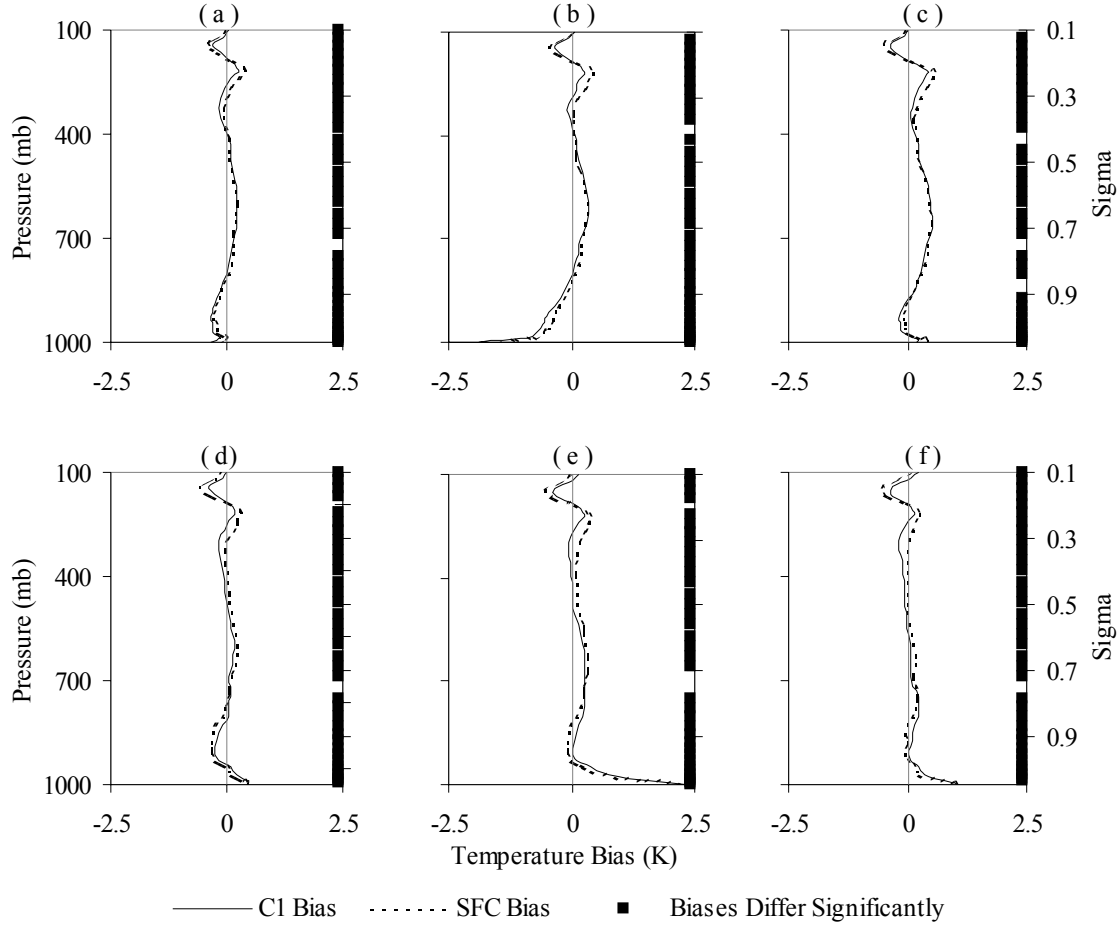
differences using the same data.

Figure 16. Average upper air temperature bias comparison (C1 vs. UPP). Same as Figure 14.

4.2.2. Geopotential Height Verification.   As shown in the previous section, the UPP

and TOT test cases produce similar results with respect to the relationship between the

different test cases. Figures 17 through 19 present the geopotential height RMSE differences

between the C1 control group and the TOT, UPP, and SFC test cases respectively.

Similarities are evident between the temperature and geopotential height RMSE

analyses. Considering the total data denied case (Figure 17), the RMSE differences are
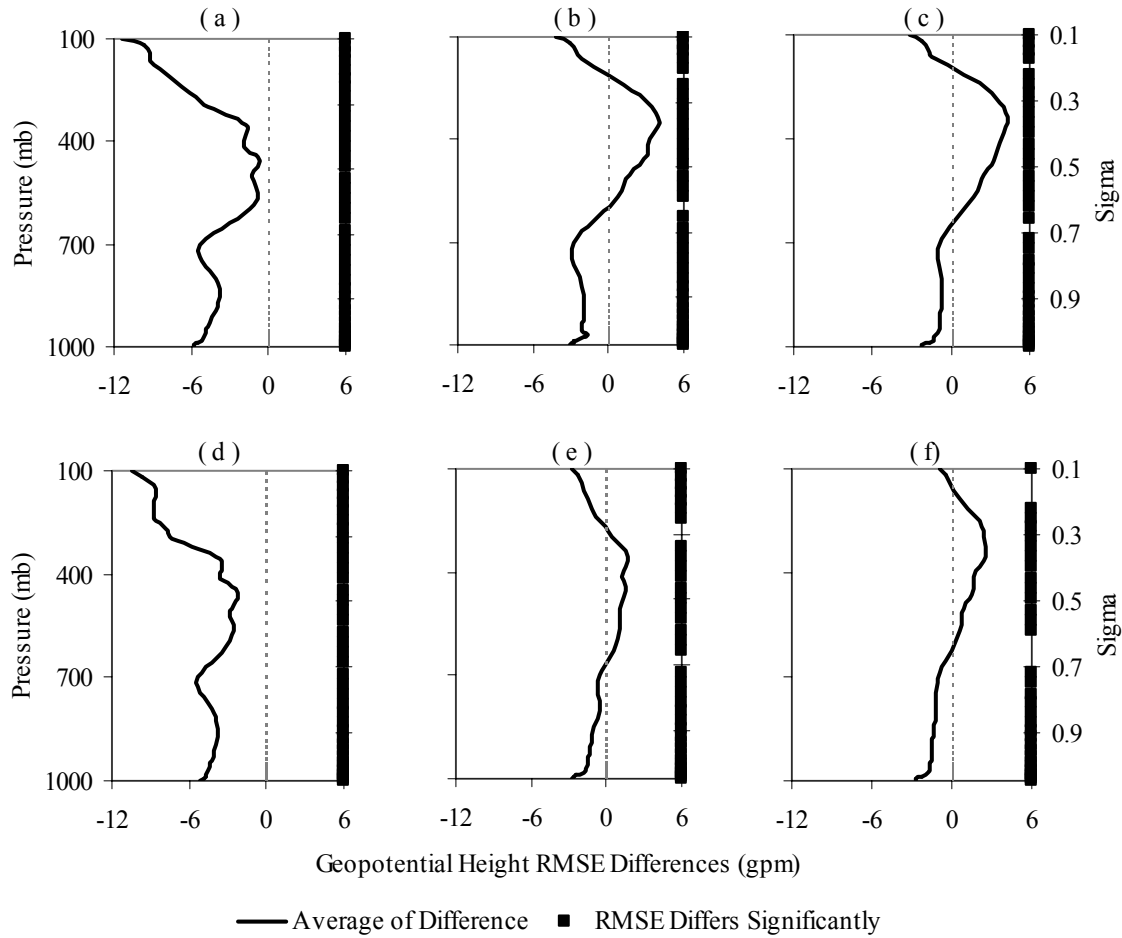
48

Figure 17. Average differences (C1-TOT) in geopotential height RMSE. Charts (a), (b), and (c) represent the 0-hr, 12-hr, and 24-hr forecasts respectively for the 00Z initialization time. Charts (d), (e), and (f) represent the 0-hr, 12-hr, and 24-hr forecasts respectively for the 12Z initialization time. The black bar denotes sigma levels with a statistically significant difference at the 95% significance level.

more significant at the initialization time than at the 12 or 24-hour forecast times. Although the trend is for the difference to decrease later in the forecast period, this is not as well defined as it is in the temperature analysis (Figure 11).

The RAOB data denial (UPP) results shown in Figure 18 indicate a similar situation as the results from the total denied case (TOT) (Figure 17). The largest differences are realized at the initialization time and these differences converge throughout the forecast period.
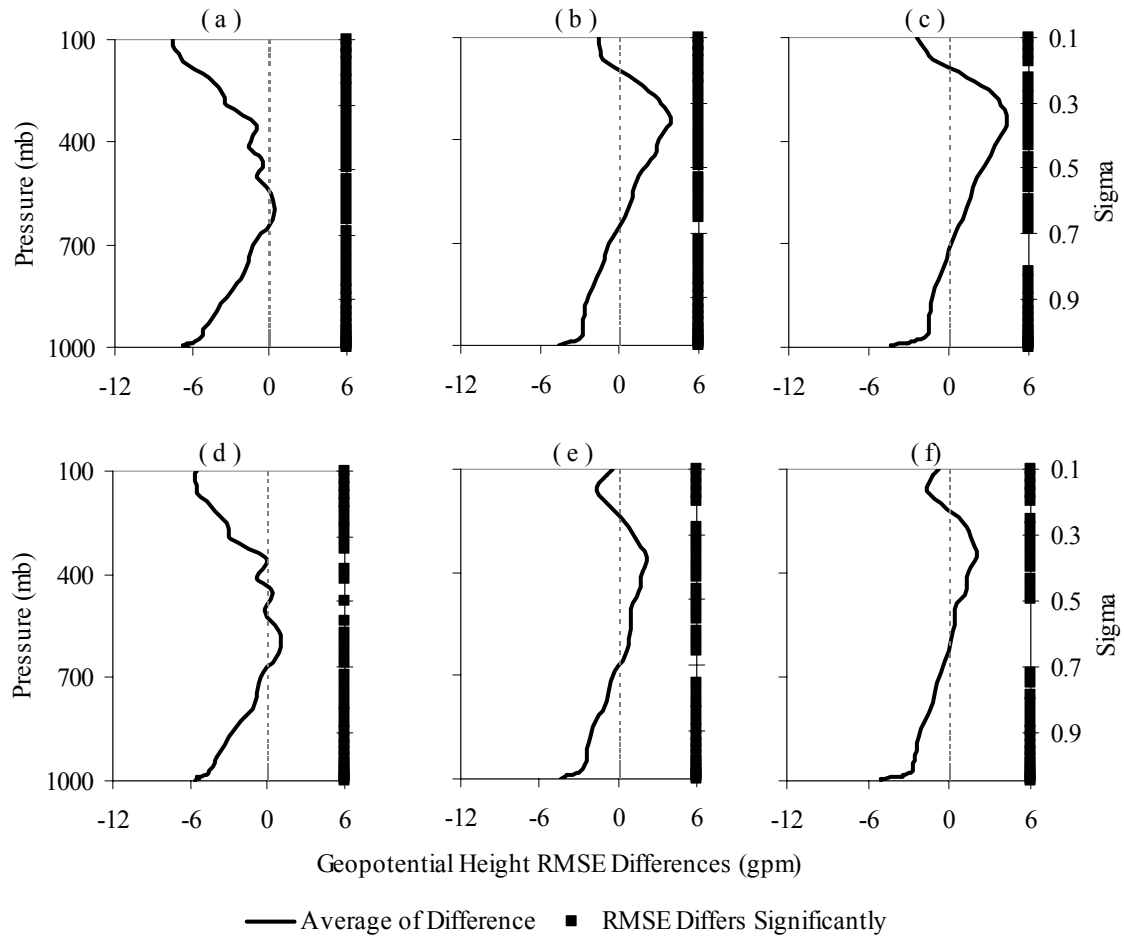
Figure 18. Average differences (C1-UPP) in geopotential height RMSE. Same as Figure 17.

Similarities are also realized in considering the RMSE differences of the vertical temperature profile and that of geopotential height for the surface data denied case (Figures 13 and 19). The findings of this study indicate that the geopotential height RMSE differences are relatively small at the model initialization time, which carries through for the entire forecast period, as with the temperature analysis. Intuitively, this trend is due to the minimal impact the surface observations have on the upper air model output.
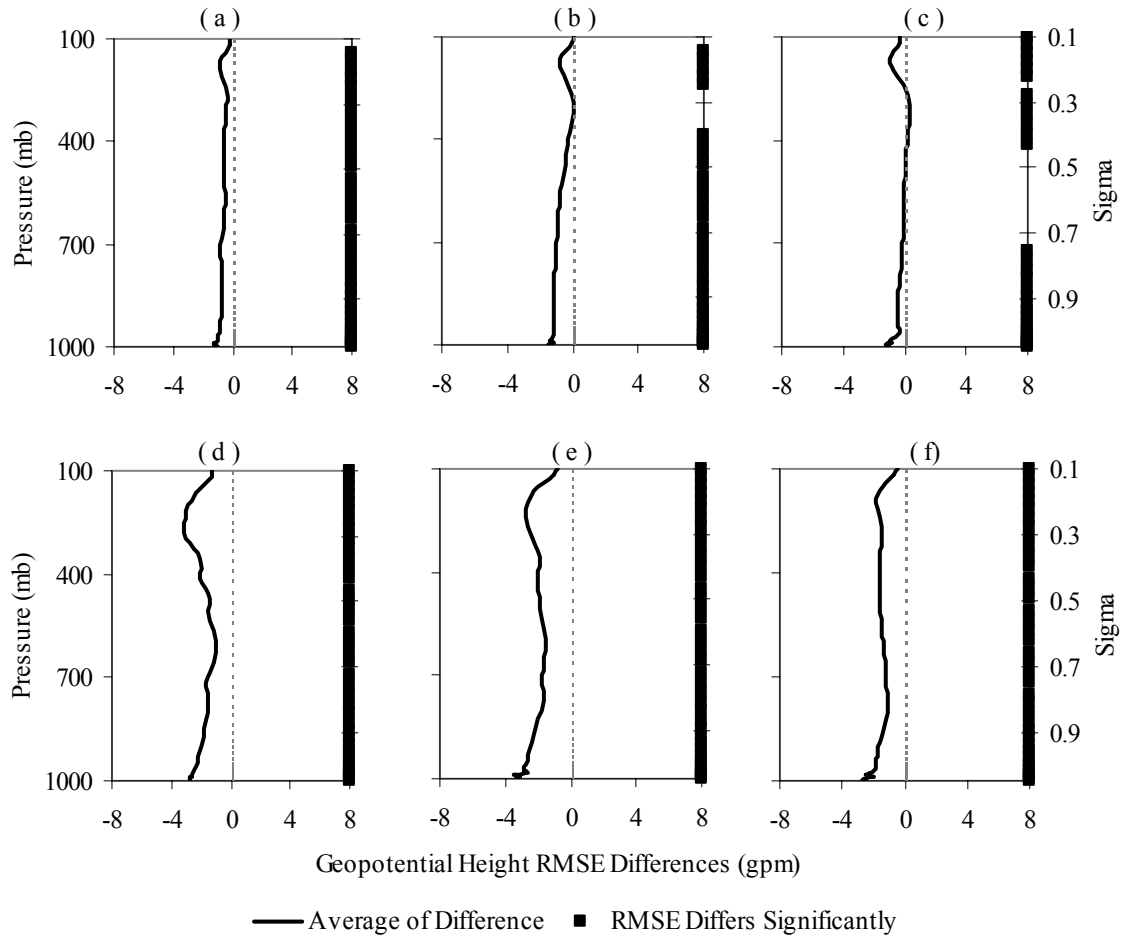
Figure 19. Average differences (C1-SFC) in geopotential height RMSE. Same as Figure 17.

The charts in Figures 20 and 21 show a comparison of the biases for the C1 control group and the UPP and SFC test cases respectively. Figure 20 shows that the biases for both cases start out at relatively similar values for both initialization times. By the 12-hour forecast, the bias values have become smaller or even negative near the surface, but have changed very little in the upper levels. The biases then become more negative near the surface by the 24-hour forecast, yet change very little in the upper levels of the atmosphere for both the 00Z and 12Z initialization times.
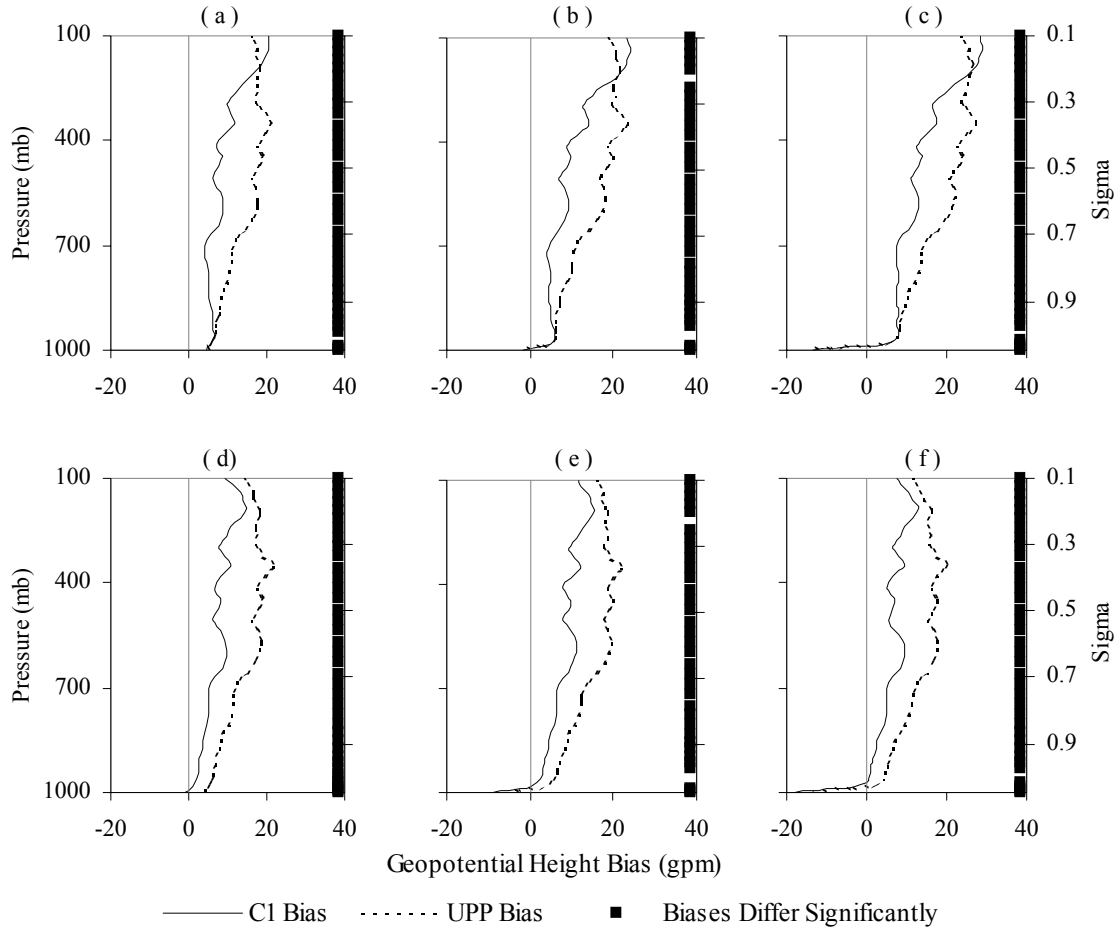
Figure 20. Average geopotential height bias comparison (C1 vs. UPP). Charts (a), (b), and (c) represent the 0-hr, 12-hr, and 24-hr forecasts respectively for the 00Z initialization time. Charts (d), (e), and (f) represent the 0-hr, 12-hr, and 24-hr forecasts respectively for the 12Z initialization time. The black bar denotes sigma levels with a statistically significant difference at the 95% significance level.

Analysis of the charts in Figure 21 indicates approximately the same outcome as that for the UPP case (Figure 20) with the exception that the SFC data are closer to C1 than those of UPP. The relative difference between C1 and SFC in the biases is carried through from the model initialization time through the 24-hour forecast.
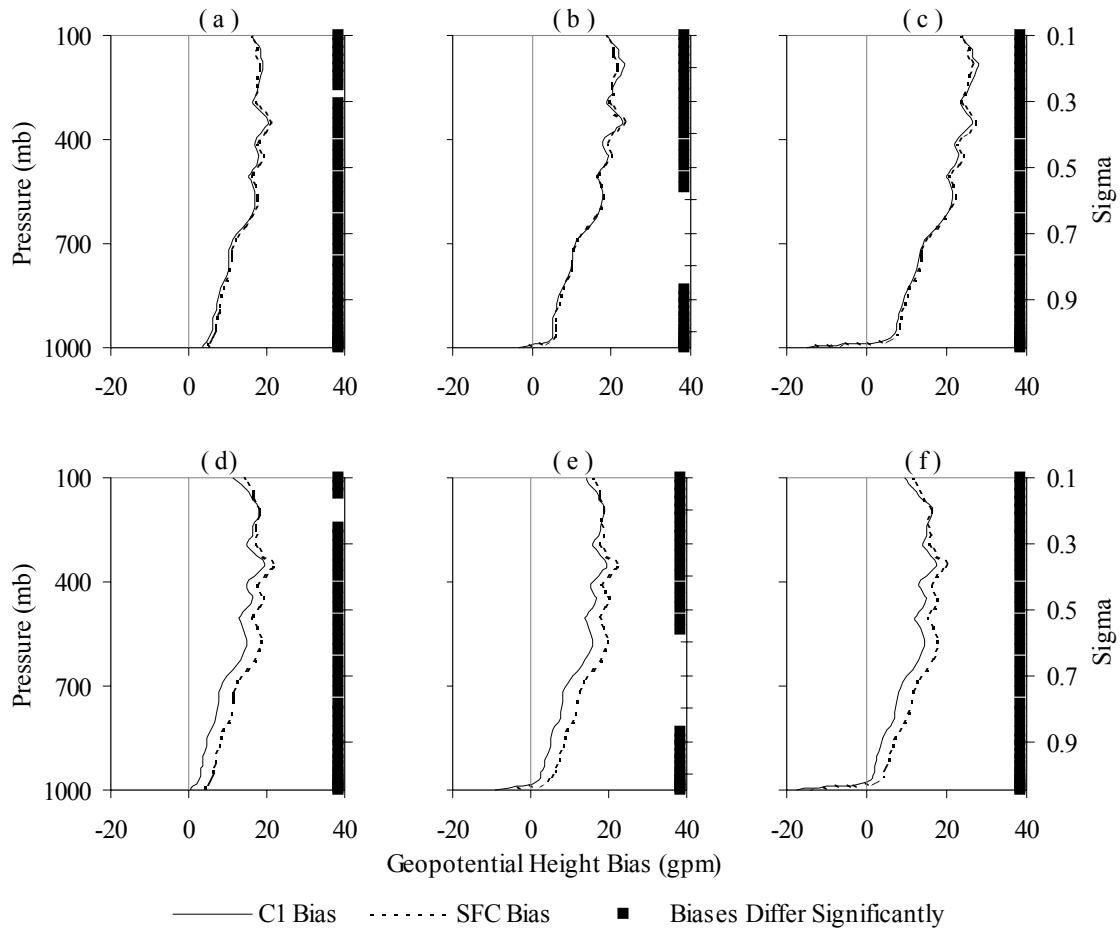
Figure 21. Average geopotential height bias comparison (C1 vs. SFC). Same as Figure 20.

4.3. Surface Verification

This section presents the findings of the surface verification process. The surface

verification statistics were calculated using the 86 surface observation stations listed in

Appendix A. The surface parameters analyzed in this study were temperature, mean sea

level pressure, and winds with all analysis in this section presented in that order. The surface

verification is accomplished for the model initialization time and at every three-hour point

thereafter up to the 24-hour forecast. Separate analyses are presented for the three different

test cases, as they were for the upper air verification. The same paired *t*-tests, including null hypothesis and confidence interval used for the upper air verification are used for all surface verifications.

As with the upper air verification, the black bars in the surface charts show the results of the paired t-test. The presence of a black bar indicates that there is a statistically significant difference between the compared model outputs at that point. The absence of a bar suggests that the differences between the parameters were not large enough to be statistically significant.

4.3.1. Temperature Verification. Figures 22 and 23 present the results of the RMSE and bias analyses for TOT test case at the 00 and 12Z initialization time respectively, Figures 24 and 25 show the same information for the UPP case, and Figures 26 and 27 depict both analyses for the SFC case. The charts labeled *(a)* in each figure indicate the average RMSE over the entire domain for the specific test case, while the charts labeled *(b)* indicate the average biases.

The gradual increase in RMSE throughout the entire forecast period evident in all of the average RMSE charts is to be expected. This trend is due to the fact that errors in the model output increase as the model gets further away from its initialization time. Another feature of note is the diurnal variation in the temperature bias as indicated by the dip (or arch) in the trace. This characteristic is normal in the MM5 output and is not indicative of an error created by the data denial processes of this study. In the European theater, 00Z corresponds to midnight local and 12Z corresponds to noon local. These are the times of minimum and maximum heating, thus leading to diurnal surface temperature variations. The MM5 does
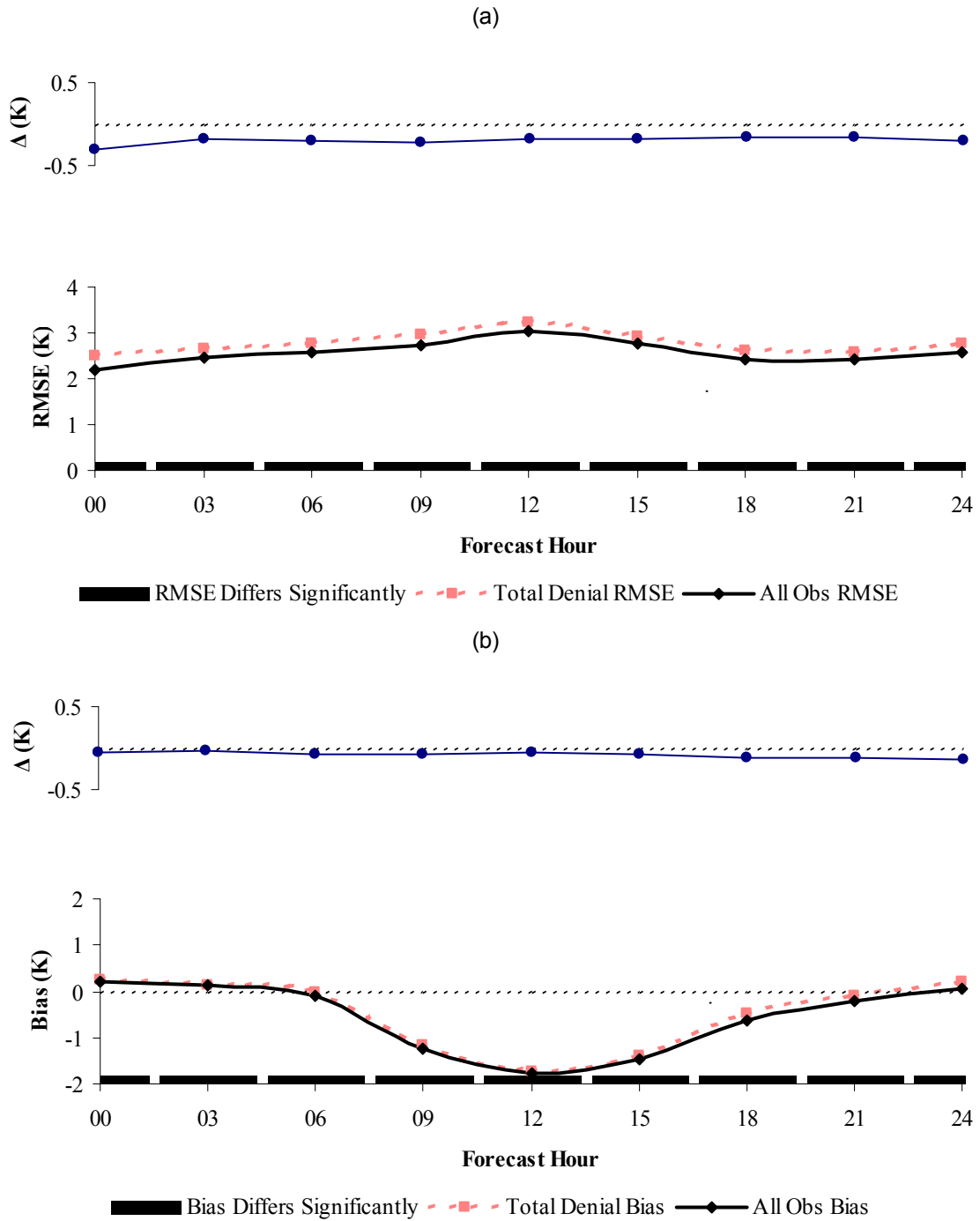
(a)

(b)

Figure 22. 00Z initialization surface temperature verification results (TOT). The graphs of Δ represent the average differences (C1-TOT) as previously defined. Chart (a) depicts the average RMSE and chart (b) is the average bias.
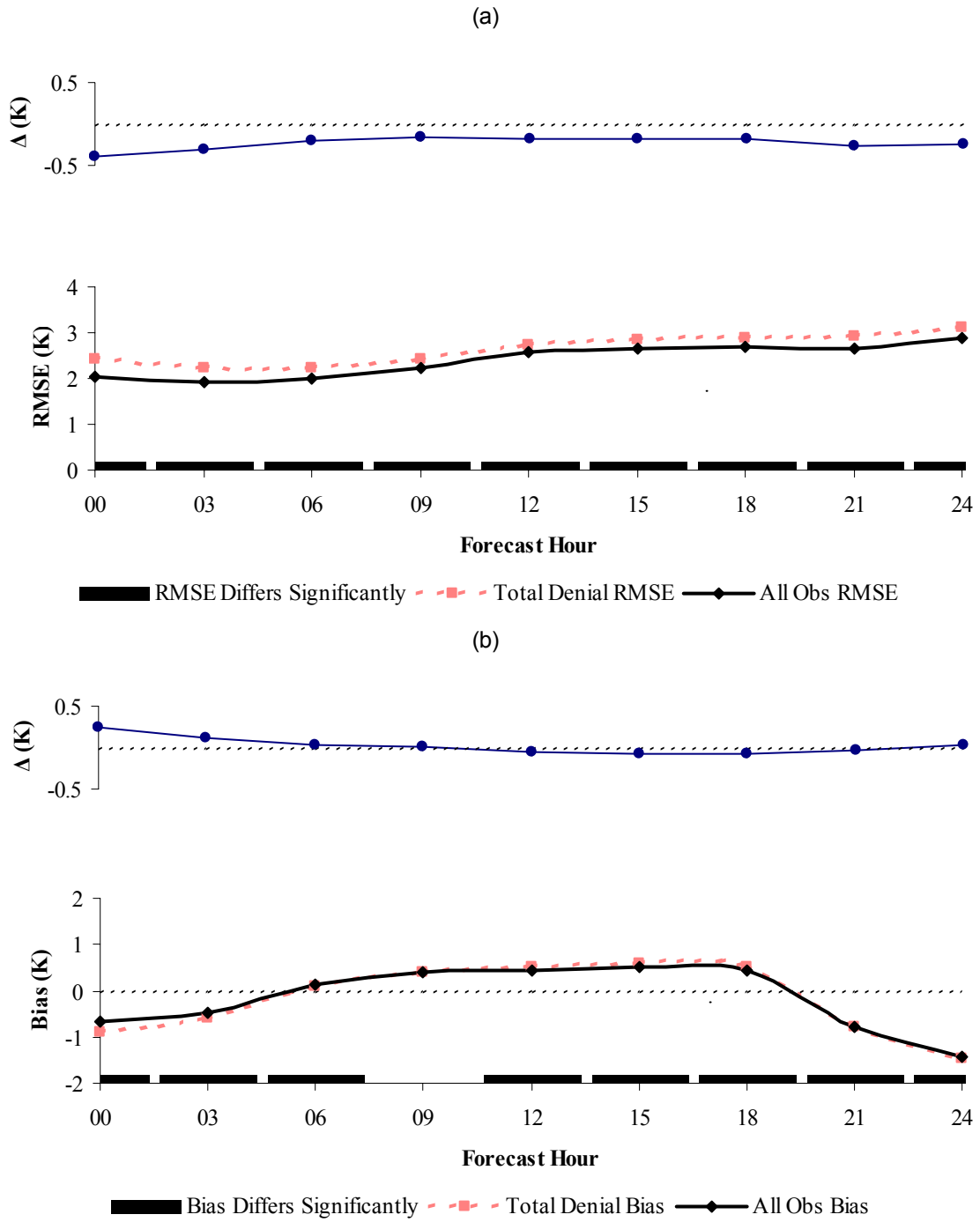
(a)

(b)

Figure 23. 12Z initialization surface temperature verification results (TOT). The graphs of Δ represent the average differences (C1-TOT). Chart (a) depicts the average RMSE and chart (b) is the average bias.

Figure 24. 00Z initialization surface temperature verification results (UPP). The graphs of Δ represent the average differences (C1-UPP). Chart (a) depicts the average RMSE and chart (b) is the average bias.
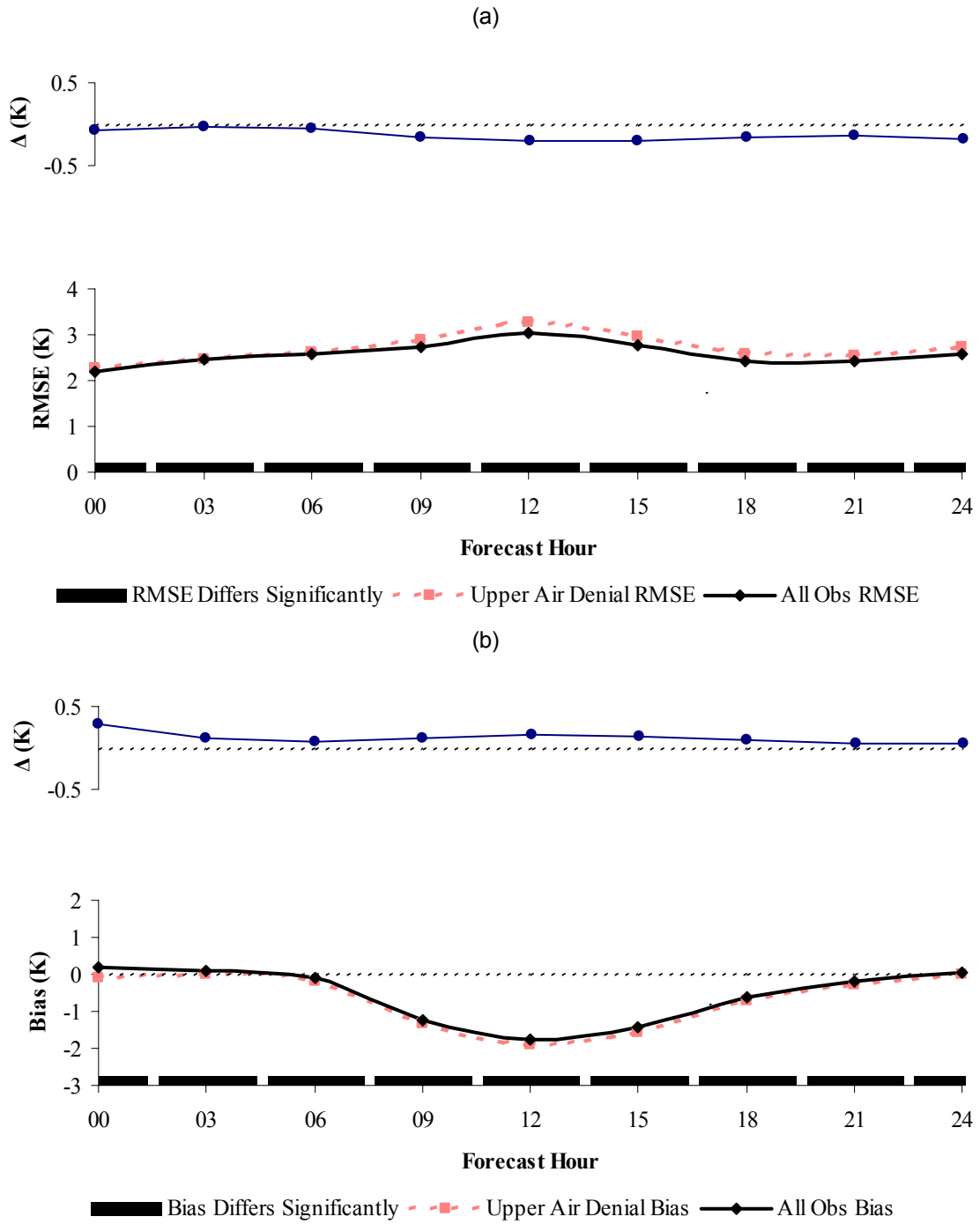
(a)

(b)
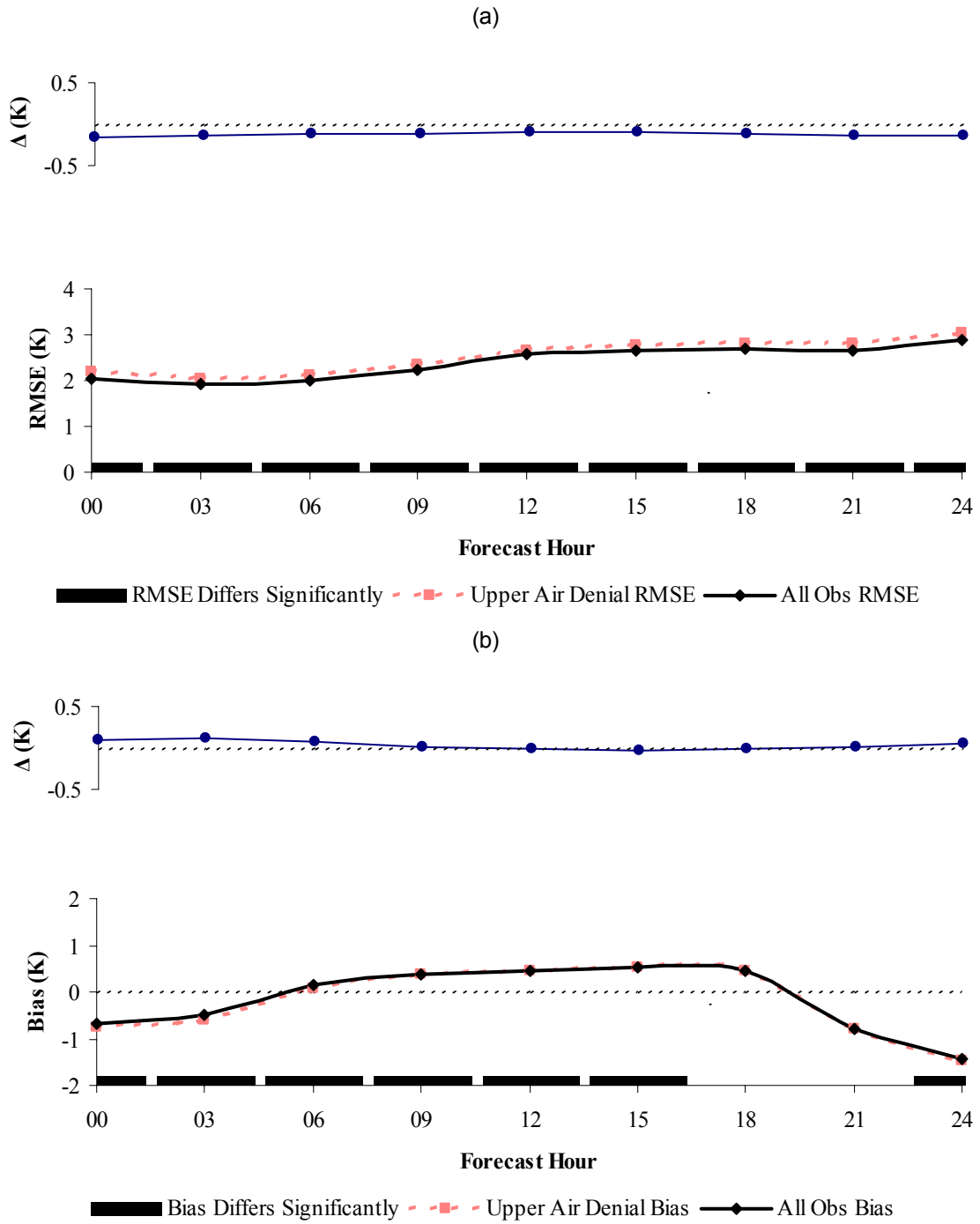
Figure 25. 12Z initialization surface temperature verification results (UPP). The graphs of Δ represent the average differences (C1-UPP). Chart (a) depicts the average RMSE and chart (b) is the average bias.

(a)

(b)

Figure 26. 00Z initialization surface temperature verification results (SFC). The graphs of Δ represent the average differences (C1-SFC). Chart (a) depicts the average RMSE and chart (b) is the average bias.

(a)

(b)

Figure 27. 12Z initialization surface temperature verification results (SFC). The graphs of Δ represent the average differences (C1-SFC). Chart (a) depicts the average RMSE and chart (b) is the average bias.
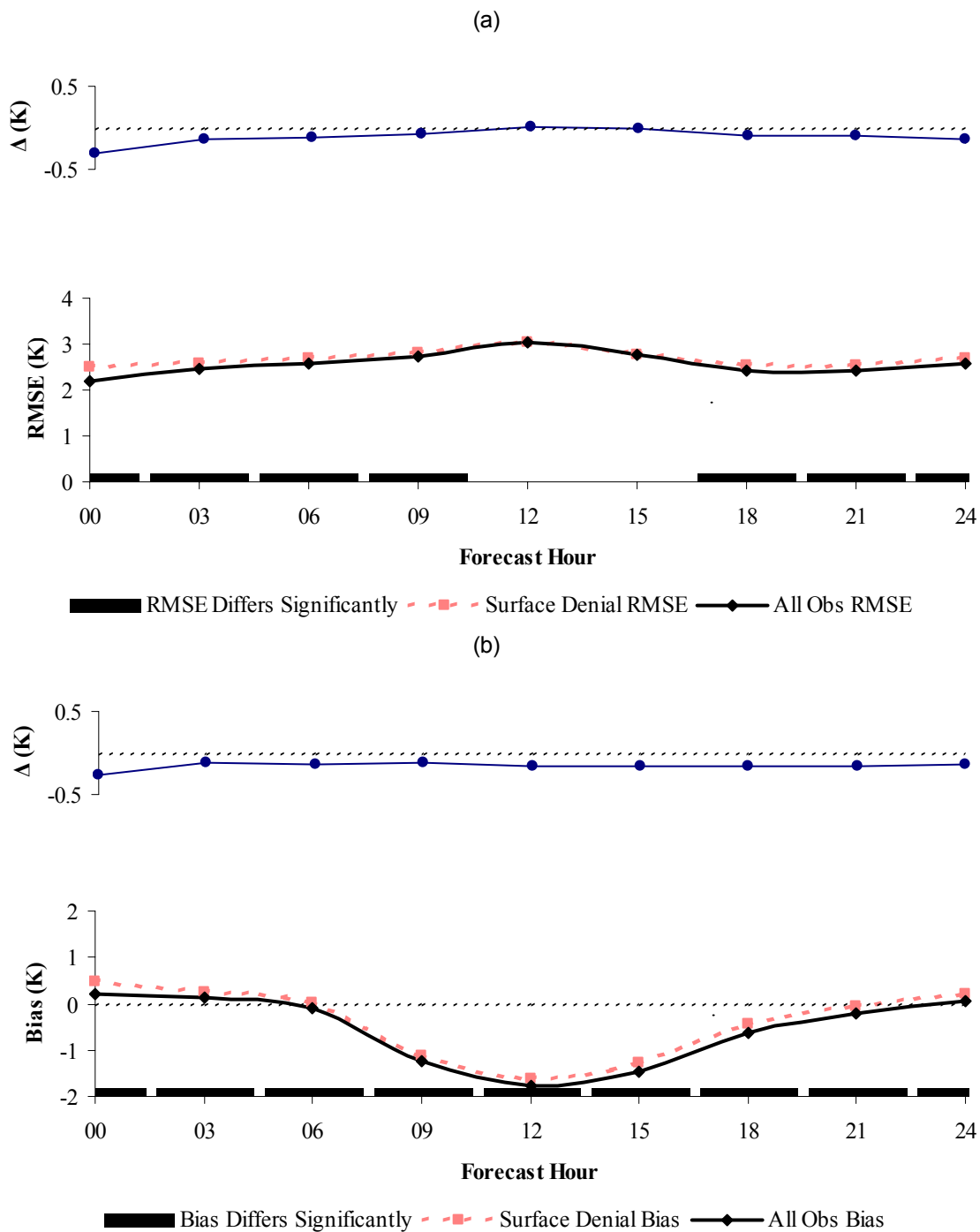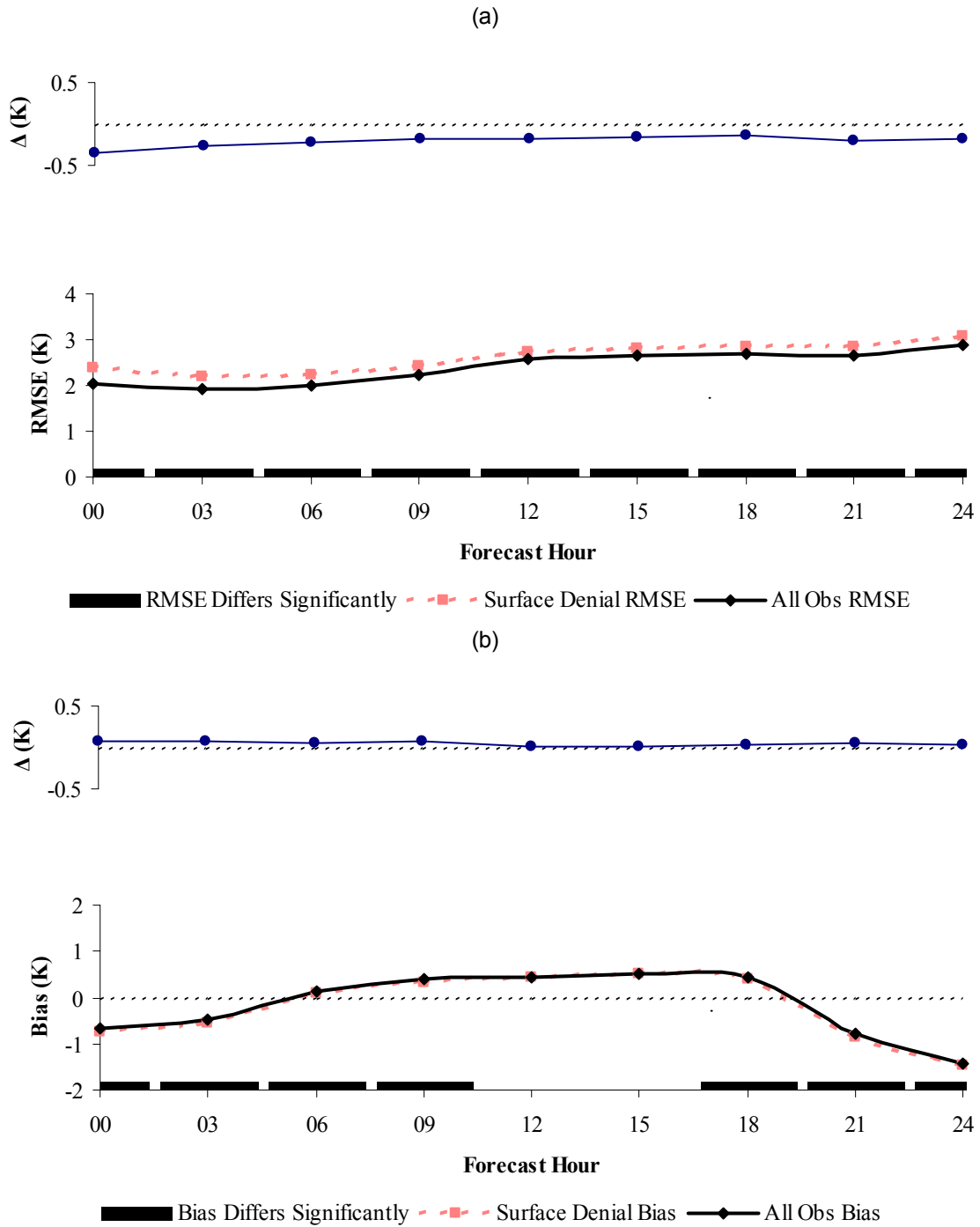
not respond well to these temperature fluctuations, thus there is a negative bias in the model output at the peak of daytime heating and a positive bias around 00Z.

The diurnal variation is also evident in the RMSE, but is not as pronounced. Although these features are noteworthy and merit discussion, neither is the principal focus of this study. The primary focus here is to analyze the difference in the test cases as compared to the control group, as the control group is assumed to be representative of the model forecast as produced by the AFWA production MM5 using all available observations.

In all instances, the average RMSE charts indicate that the C1 control group initialized better than the data denied test case. The different test configurations have varying levels of impact on the output of their respective runs. The TOT test case produced the largest RMSE differences, which is to be expected since this setup allowed no current observations to be assimilated. The analyses of the UPP and SFC test cases show similar results, however not as pronounced. All of the differences remain relatively constant throughout the entire forecast period. All cases indicate the difference to be less than 0.4 K for both RMSE and bias, with the majority of them being less than 0.2 K, which is well within the accuracy of the surface temperature measuring devices as listed in Table 6.

The bias charts reveal similar results for all of the data denied test cases. For both initialization times and all test cases, the differences are small and remain relatively constant throughout the entire forecast period. Although the data were determined to be statistically different, the differences are not meteorologically significant in most cases because the errors are less than the possible error associated with the equipment used to take the observations.

4.3.2. Mean Sea Level Pressure Verification. The observed altimeter settings as reported by the surface verification stations were used as a basis of comparison for mean sea level pressure. Since the model does not produce any actual surface output, the pressure calculated at the lowest sigma level was reduced to the surface as described in section 3.5.3. This extrapolated model value was then converted to sea level pressure in the same manner as the observed altimeter setting to ensure consistency.

The figures in this section are presented in the same manner as those for the temperature verification. Figures 28 and 29 show the TOT case, Figures 30 and 31 show the UPP case, and Figures 32 and 33 depict the SFC case. The trends in the analyses are similar to those of the temperature verification with one minor departure; there is no specific diurnal pattern to the bias traces.

The error comparisons for pressure are similar to those found in the temperature verification. The RMSE values for each of the data denied test cases were greater than those of the C1 control group with all differences being less than 0.8 mb and the vast majority less than 0.5 mb, which is the accuracy of the sensor. The overall tendency for the average RMSE comparisons was that the difference remained relatively constant throughout the forecast period.

The significant finding in the bias analyses was that they had the same trend as RMSE, but the difference was not as consistent as that of the average RMSE comparisons. This result is a deviation from the findings of the temperature verification.

(a)

(b)

Figure 28. 00Z initialization mean sea level pressure verification results (TOT). The graphs of Δ represent the average differences (C1-TOT). Chart (a) depicts the average RMSE and chart (b) is the average bias.

Figure 29. 12Z initialization mean sea level pressure verification results (TOT). The graphs of Δ represent the average differences (C1-TOT). Chart (a) depicts the average RMSE and chart (b) is the average bias.
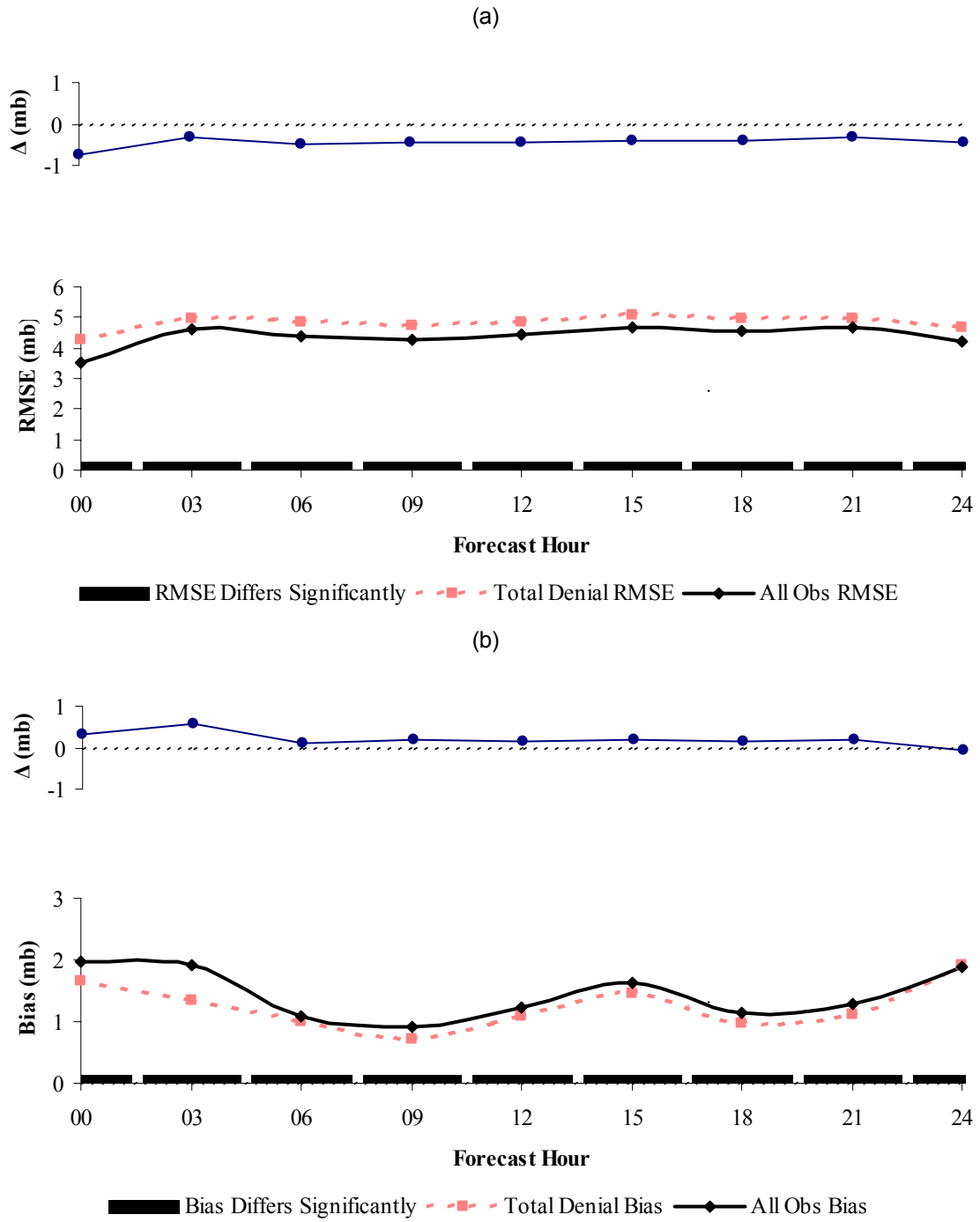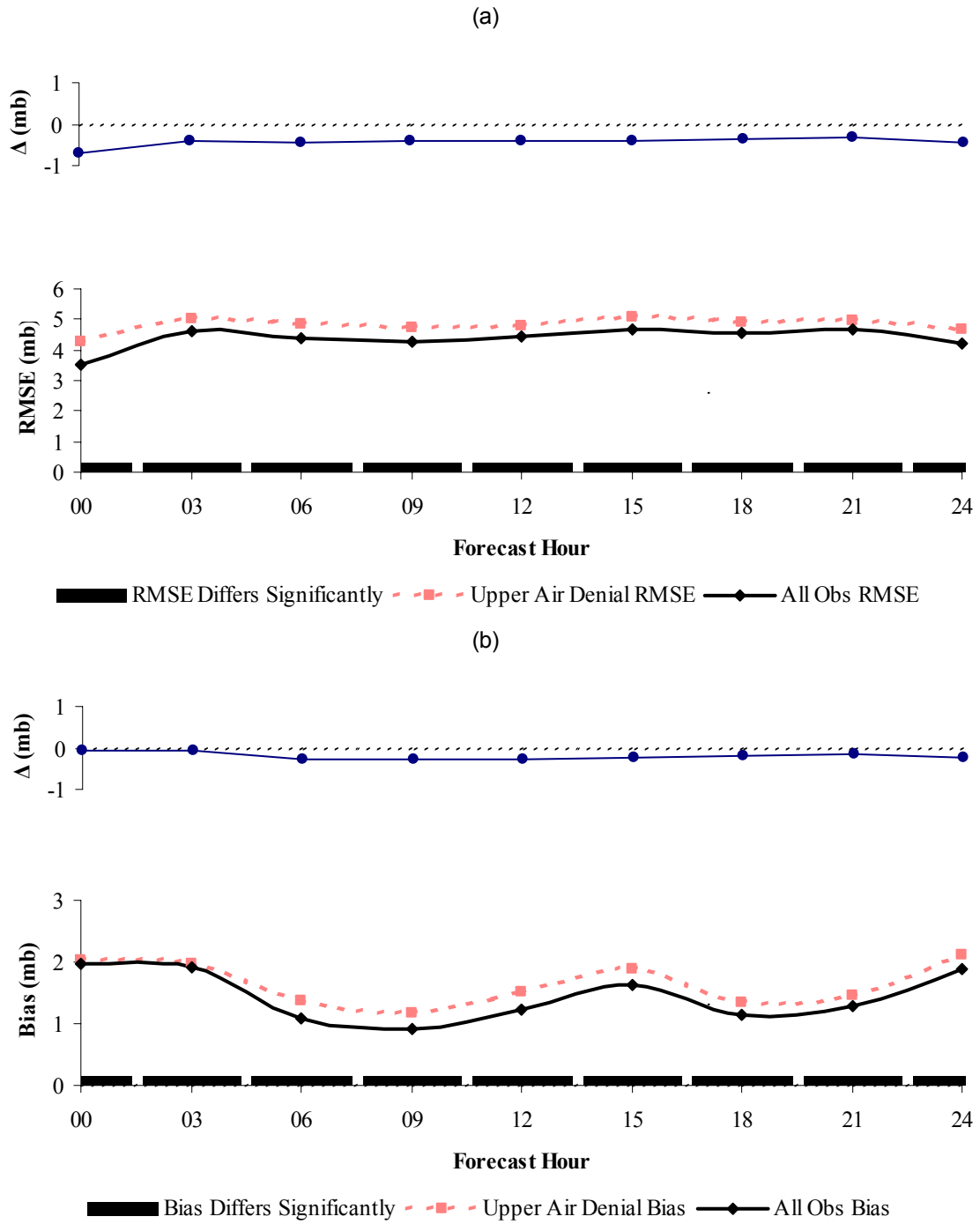
Figure 30. 00Z initialization mean sea level pressure verification results (UPP). The graphs of Δ represent the average differences (C1-UPP). Chart (a) depicts the average RMSE and chart (b) is the average bias.

(a)

(b)
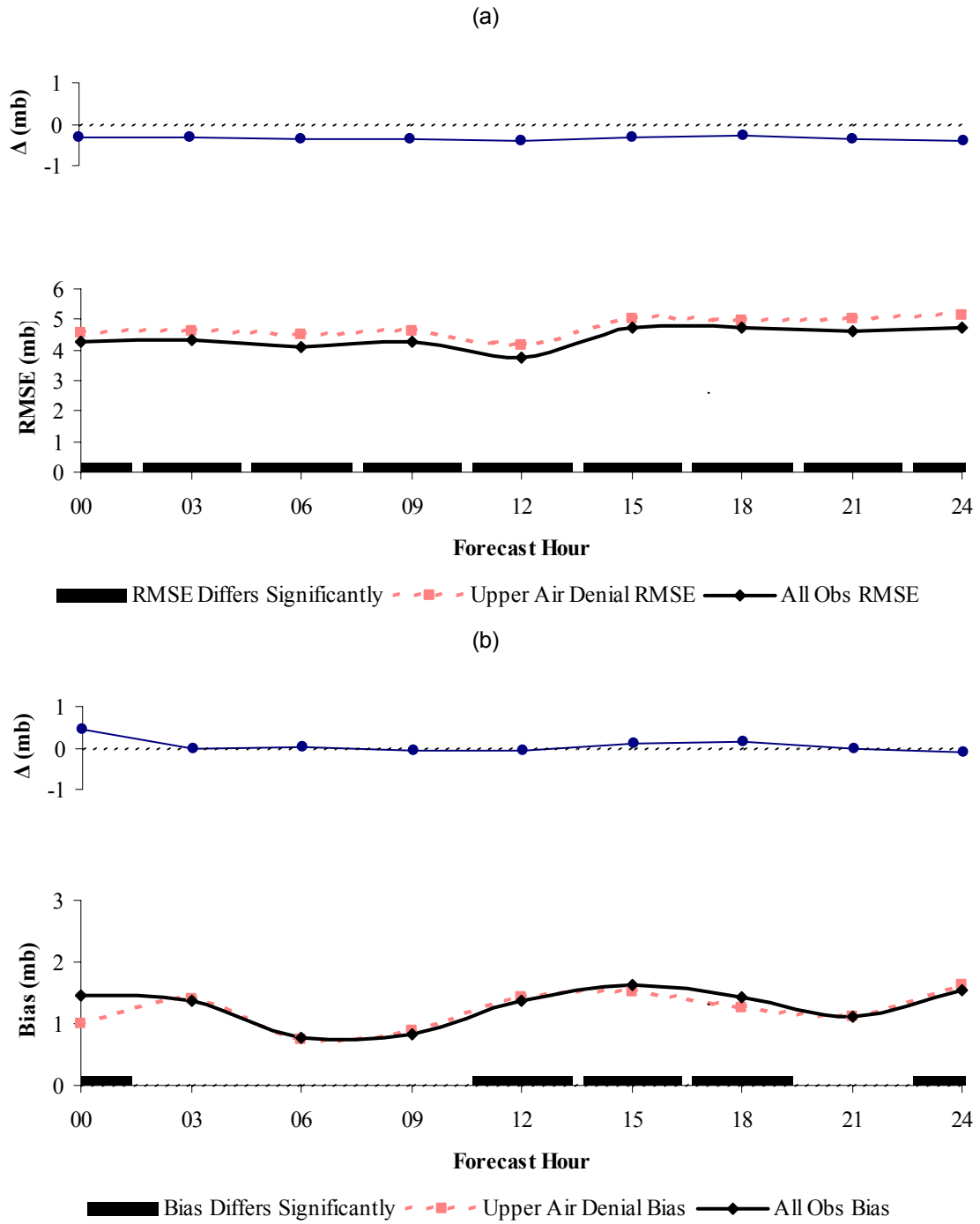
Figure 31. 12Z initialization mean sea level pressure verification results (UPP). The graphs of Δ represent the average differences (C1-UPP). Chart (a) depicts the average RMSE and chart (b) is the average bias.
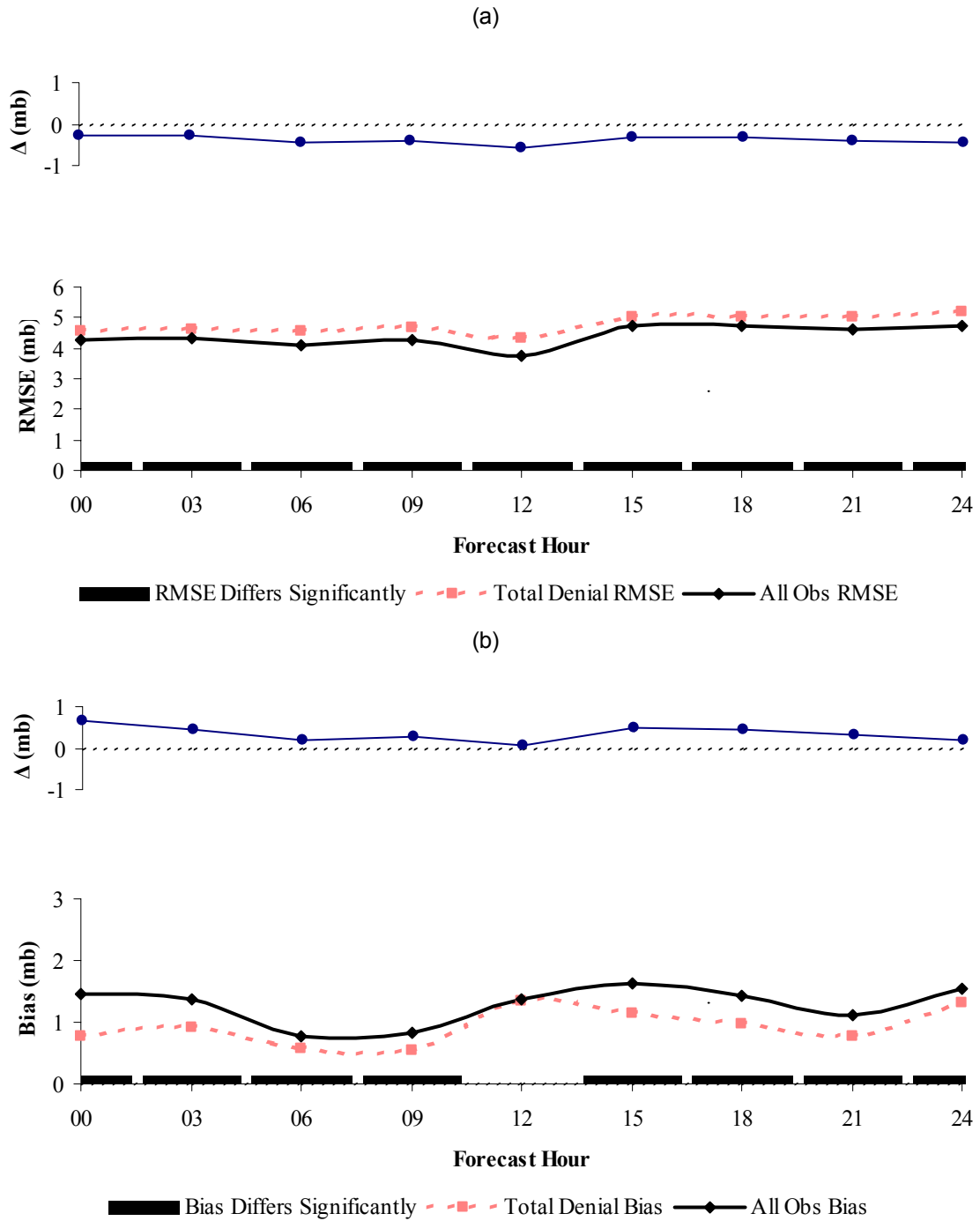
66

(a)

(b)

Figure 32.   00Z initialization mean sea level pressure verification results (SFC). The graphs of Δ represent the average differences (C1-SFC). Chart (a) depicts the average RMSE and chart (b) is the average bias.

(a)

(b)

Figure 33. 12Z initialization mean sea level pressure verification results (SFC). The graphs of Δ represent the average differences (C1-SFC). Chart (a) depicts the average RMSE and chart (b) is the average bias.
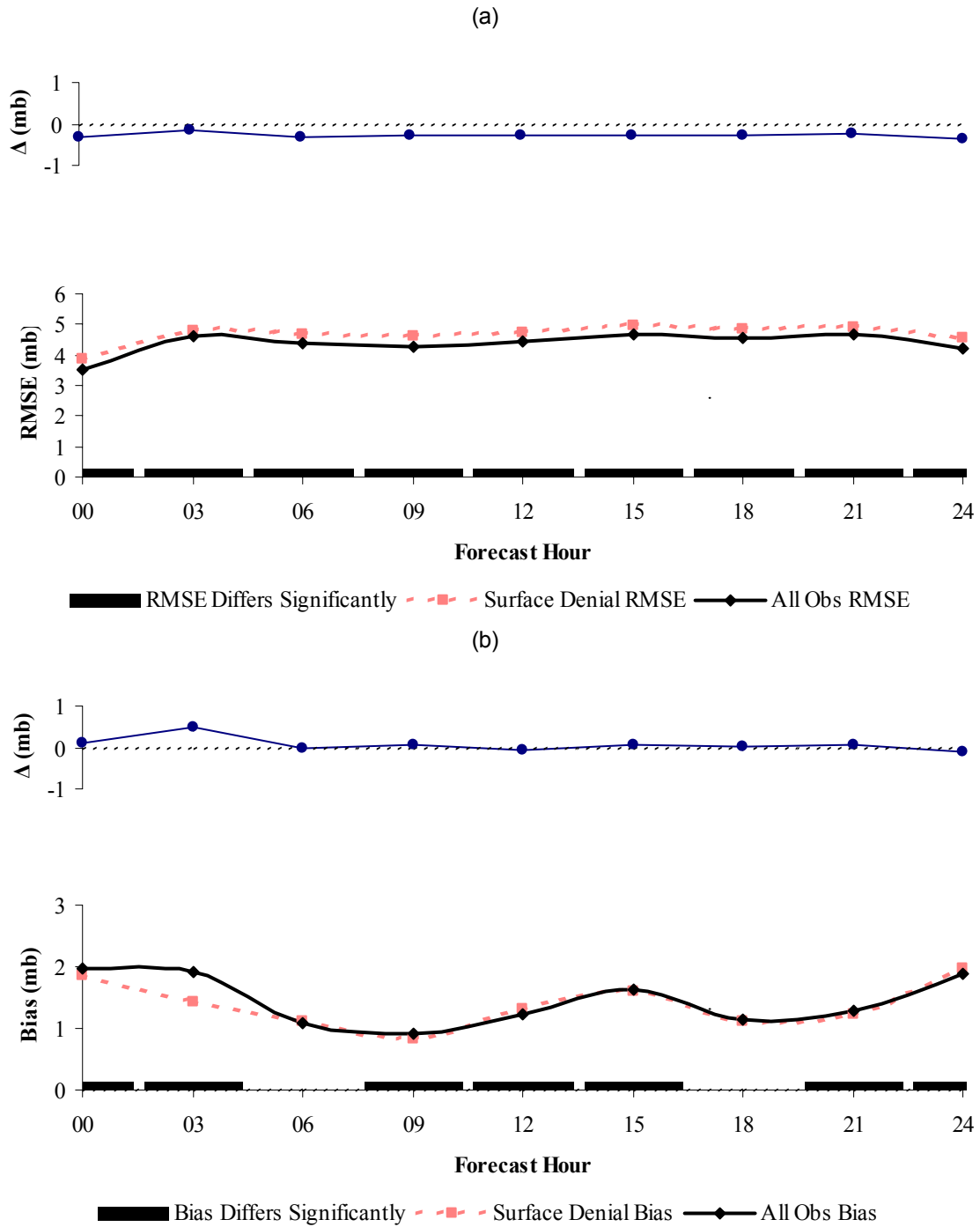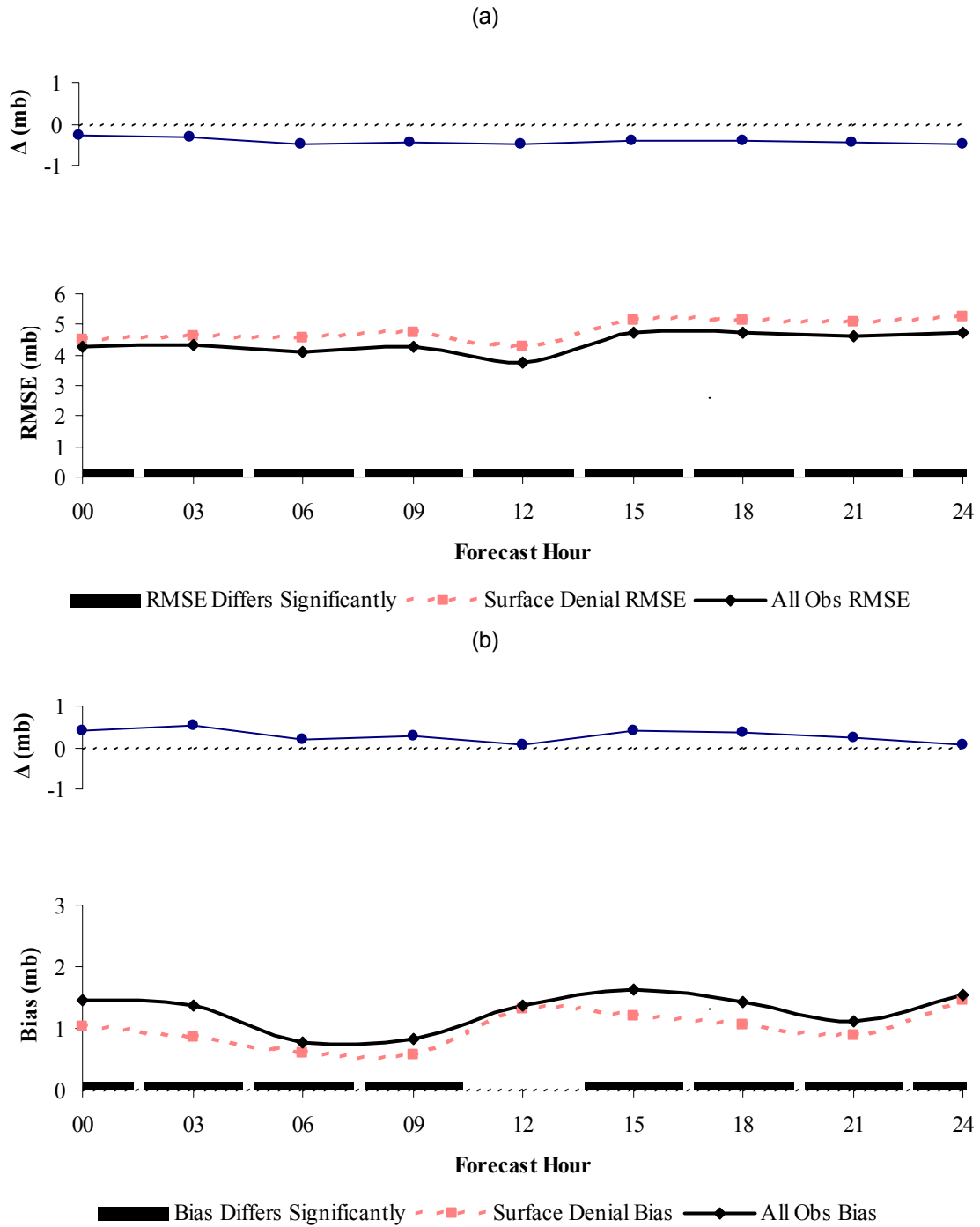
4.3.3. Surface Wind Verification.   As described in section 3.5.2, equation 5, RMSVE is the statistical tool used for wind verification. Large RMSVE values may be a direct result of erroneous wind direction predictions and may not be indicative of bad wind speed forecasts.

Figures 34 through 36 depict the average RMSVE for all data denied test cases and initialization times for three-hourly forecasts compared to the C1 control group as with the previous surface verifications. Analysis of the charts shows essentially the same trend for all test cases and both initialization times. In all instances, the tendency is for an increase in error throughout the forecast period with the differences remaining relatively constant after the 3-hour forecast time. The different test cases and initialization times produced varied results. Results from the TOT case, shown in Figure 34, indicate the average RMSVE difference at both initialization times to be slightly less than 0.9 ms^{-1}, which is greater than the accuracy of the wind measurement equipment. By the 3-hour forecast, the difference is reduced to less than 0.3 ms^{-1} where it remains relatively constant for the remainder of the forecast period. When analyzing the results of the UPP test case in Figure 37, it is evident that denial of the RAOB has a much smaller effect on the model output. Figure 36 shows the results of the SFC test case, which are similar to the results of the TOT test case. The largest difference is at the model initialization time with a sharp decrease by the 3-hour forecast time.
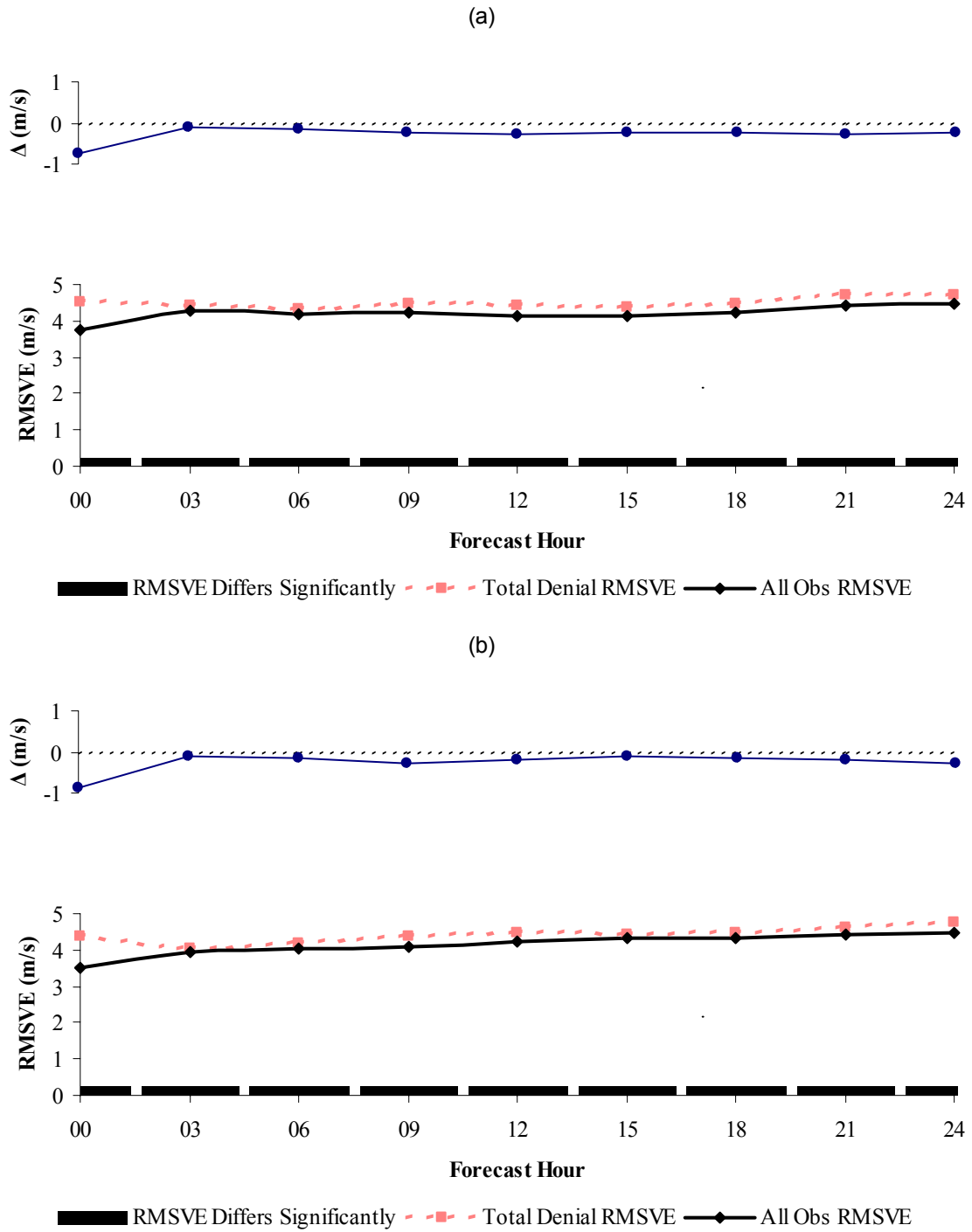
(a)

(b)

Figure 34. Surface wind speed verification results (TOT). (a) 00Z initialization. (b) 12Z initialization.
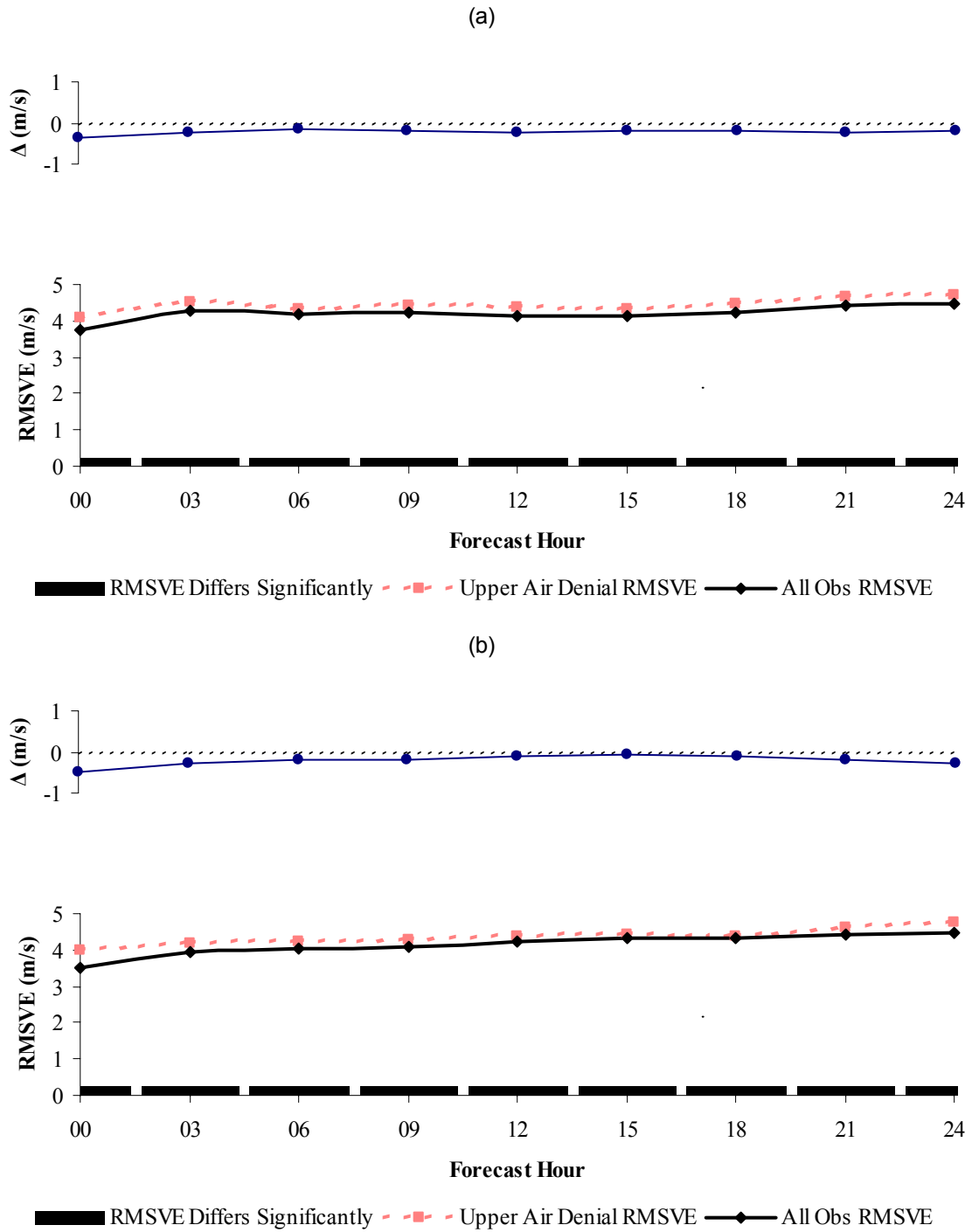
(a)

(b)

Figure 35. Surface wind speed verification results (UPP). (a) 00Z initialization. (b) 12Z initialization.
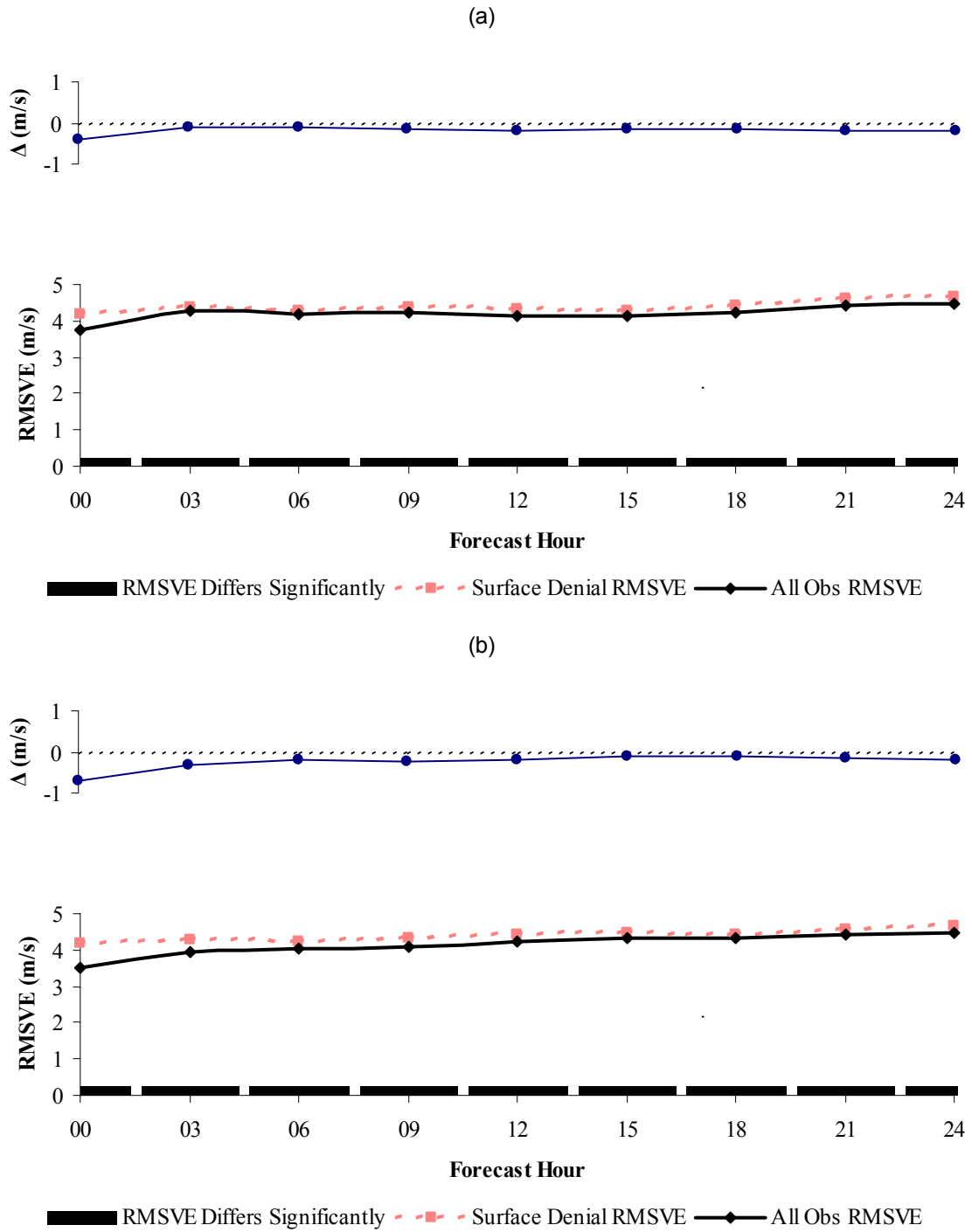
(a)

(b)

Figure 36. Surface wind speed verification results (SFC). (a) 00Z initialization. (b) 12Z initialization.

4.4. Errors due to Initial and Boundary Conditions

This section provides a comparison of the previously discussed analyses using the C1 control group with similar analyses made using the secondary C2 control group as described in section 3.3.3. The test method incorporated was designed in such a manner as to account for differences in the initial and boundary conditions created by using different first guess forecasts for the model initializations.

All figures in this section show the RMSE and RMSVE differences of the primary C1 control run made using the 6-hour old AVN data previously discussed compared to the corresponding difference of the secondary C2 control run made using the 30-hour old AVN data with all available observations. Analyses are made for both the C1 and C2 control runs as described in section 3.5. The C2 analyses are then subtracted from the previously shown C1 analyses to calculate new RMSE and RMSVE differences. The new RMSE and RMSVE differences shown here can then be attributed primarily to the use of different first guess forecasts and provide an estimate of the amount of error associated with the different initial and boundary conditions.

4.4.1. Upper Air Analyses. Figures 37 through 39 display the temperature RMSE differences for the 00Z initialization of the TOT, UPP, and SFC test cases respectively, while figures 40 through 42 show the same data for 12Z initialization times. Panels (a) through (c) in each figure depict the RMSE differences of the C1 control group and the various data denied test cases for the 0, 12, and 24-hour forecasts respectively. Panels (d) through (f) show the RMSE differences of the C2 control group and the various data denied test cases for the 0, 12, and 24-hour forecasts respectively. Finally, panels (g) through (i) are the

differences of the two previous sets of analyses and primarily show the RMSE differences

associated with the different initial and boundary conditions.

Figure 37. Average differences (C1-C2) in temperature RMSE for the TOT 00Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-TOT test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-TOT test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.

Figure 38. Average differences (C1-C2) in temperature RMSE for the UPP 00Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-UPP test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-UPP test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.

Figure 39. Average differences (C1-C2) in temperature RMSE for the SFC 00Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-SFC test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-SFC test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.
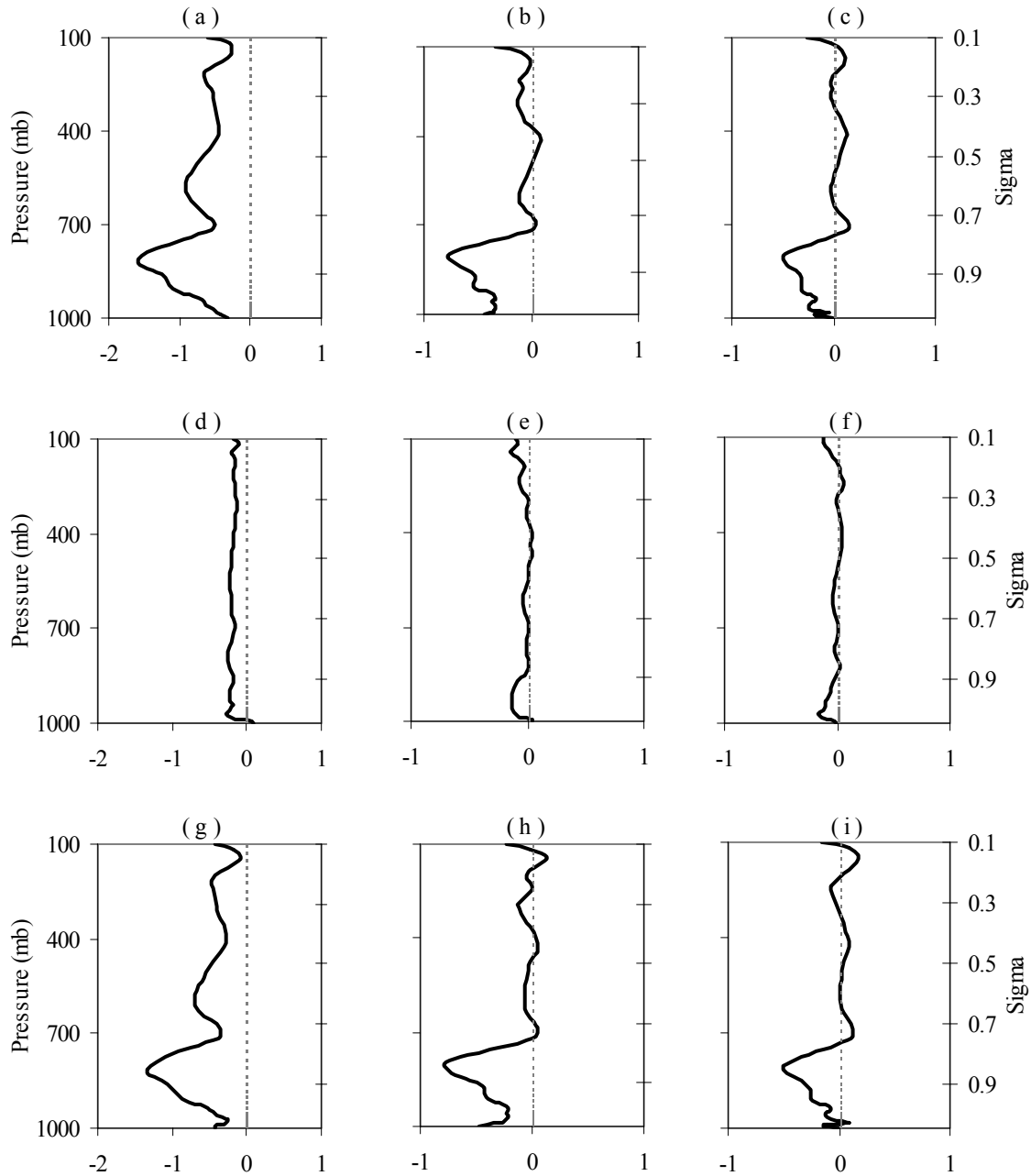
76
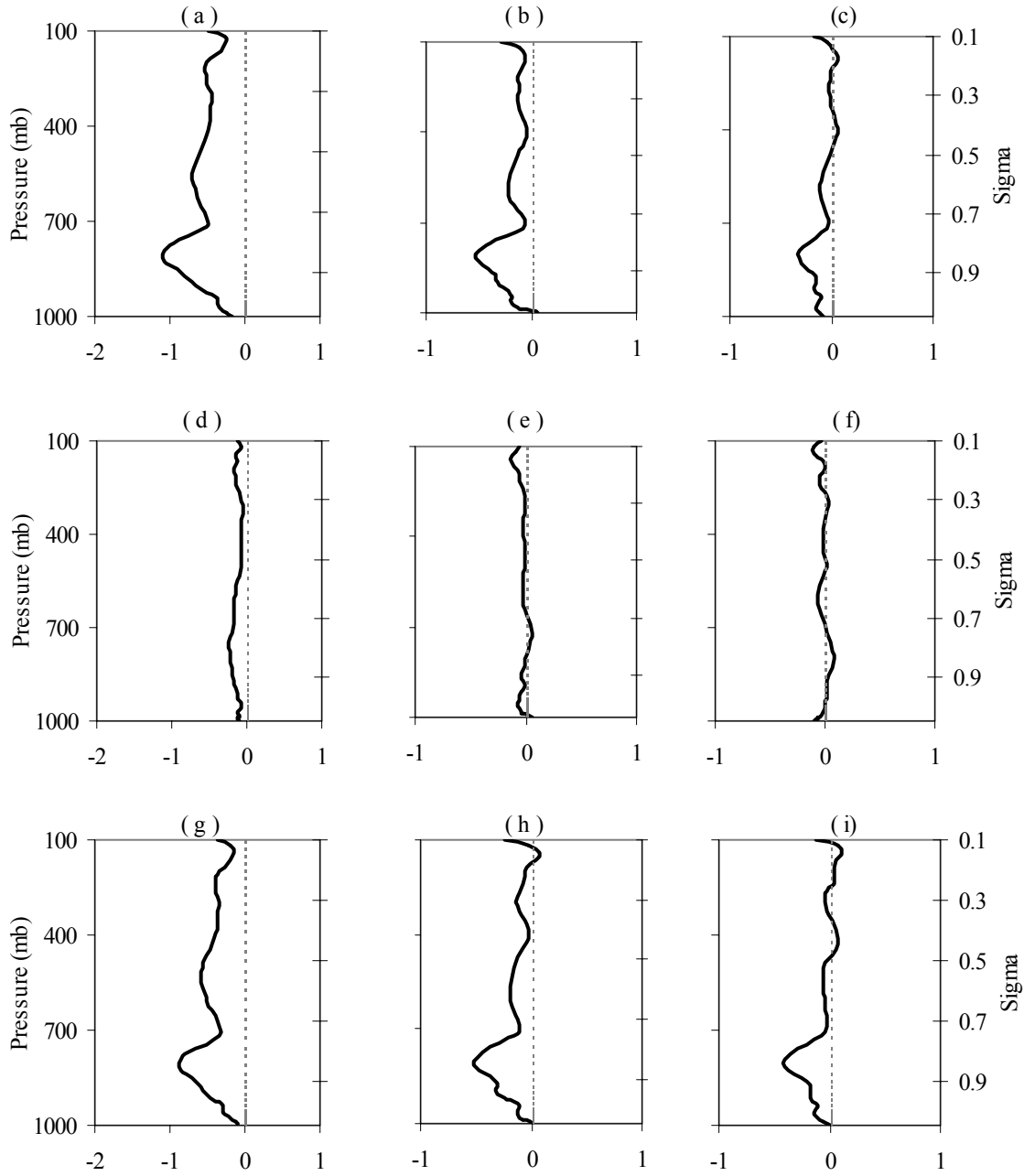
Figure 40. Average differences (C1-C2) in temperature RMSE for the TOT 12Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-TOT test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-TOT test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.
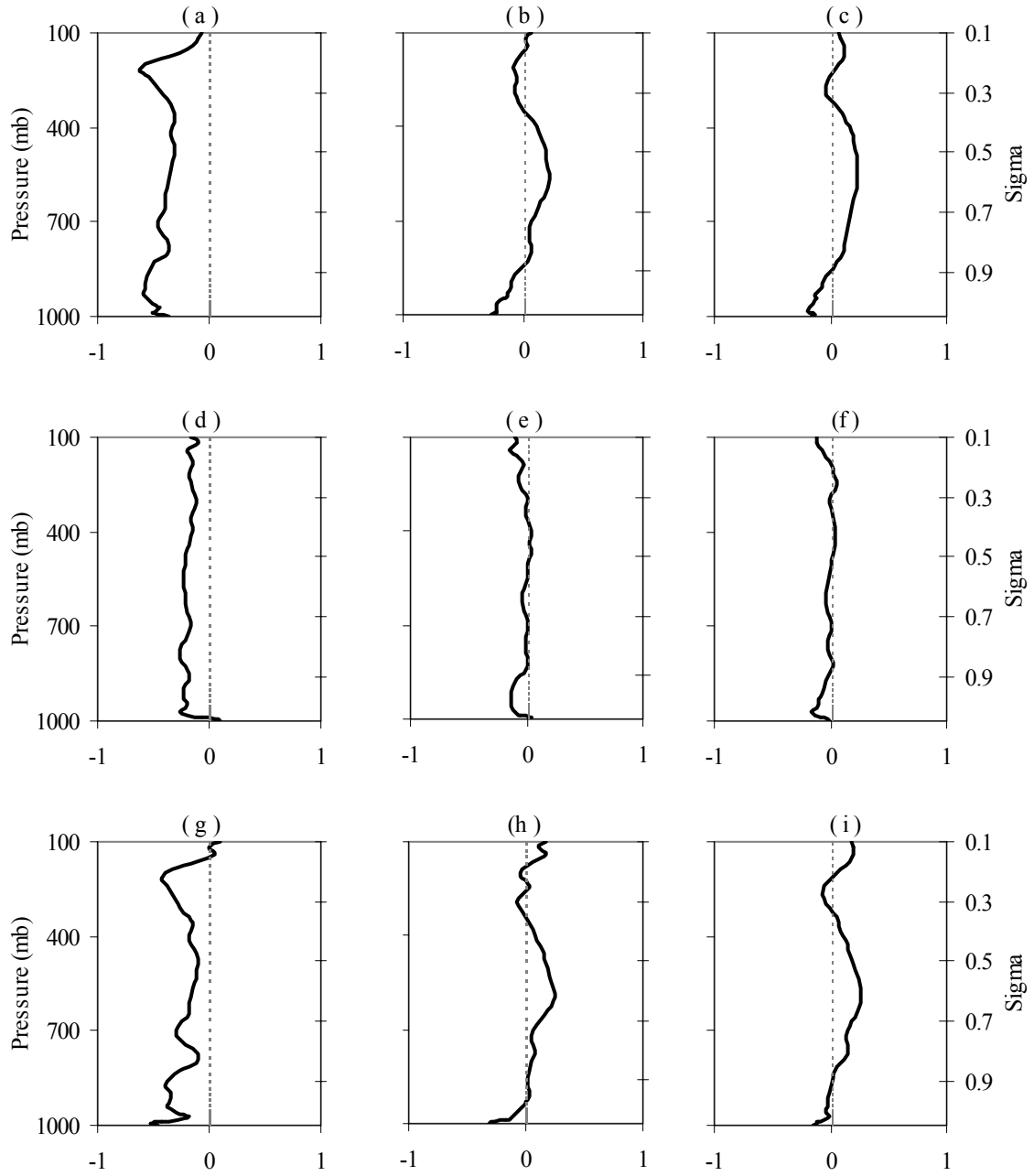
Figure 41. Average differences (C1-C2) in temperature RMSE for the UPP 12Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-UPP test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-UPP test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.
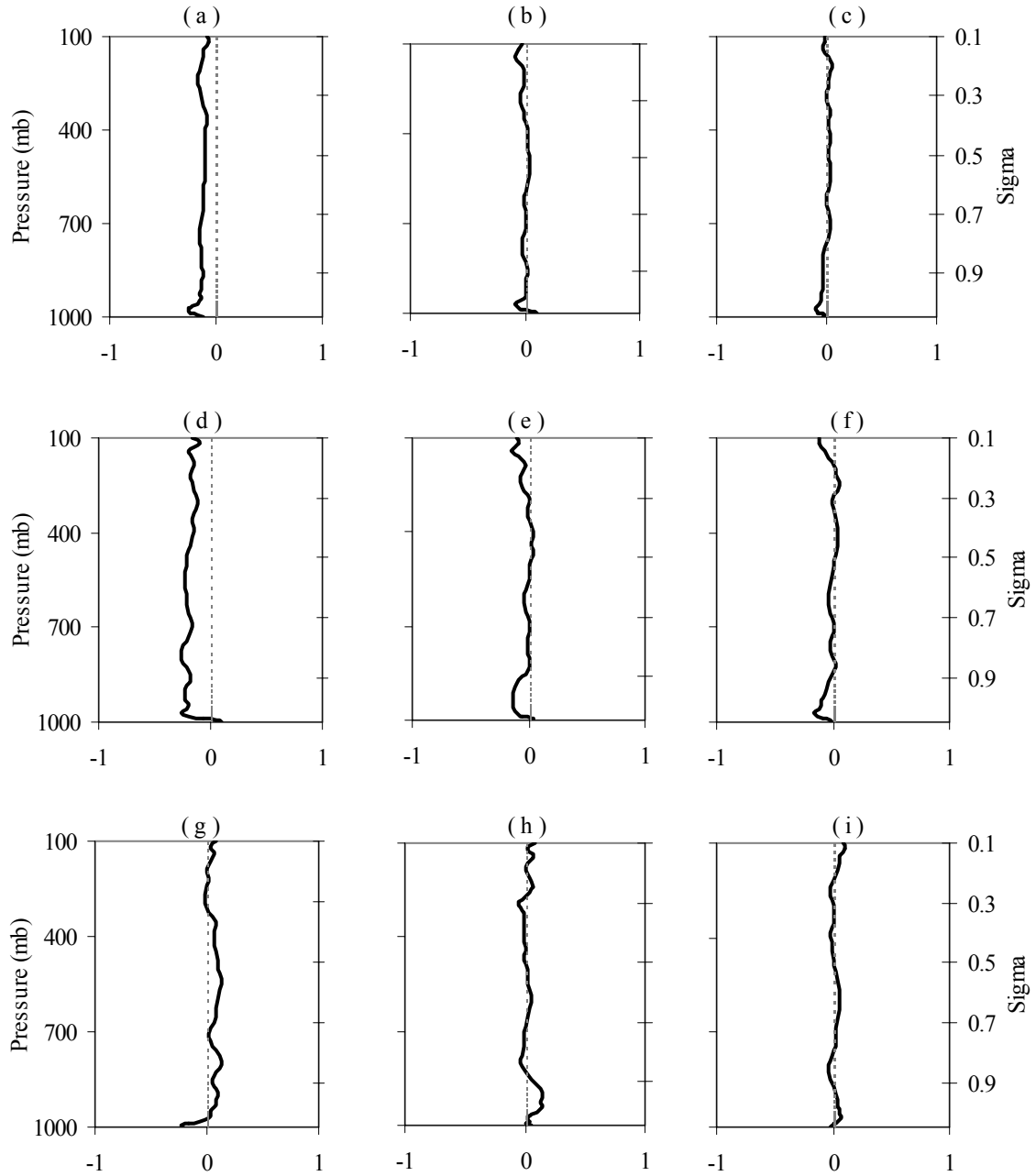
78

Figure 42. Average differences (C1-C2) in temperature RMSE for the SFC 12Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-SFC test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-SFC test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.

Figures 43 through 48 show the RMSE differences for the geopotential heights in the

exact same manner as the previous figures showed them for the upper air temperatures. The

first three figures depict the three different test cases for the 00Z initialization time, while the final three figures are for the 12Z initialization.

Figure 43. Average differences (C1-C2) in the geopotential height RMSE for the TOT 00Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-TOT test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-TOT test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.
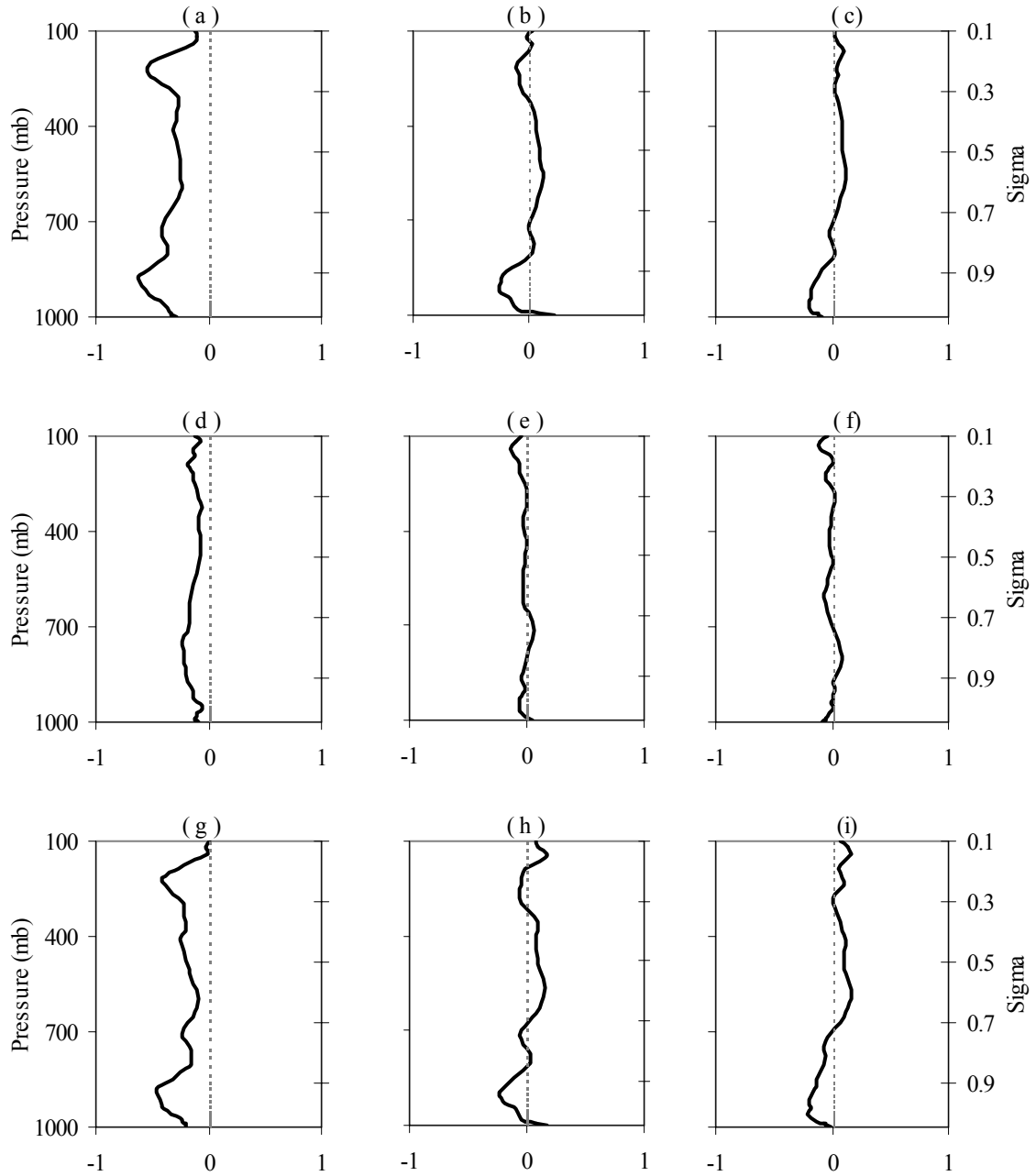
Figure 44. Average differences (C1-C2) in the geopotential height RMSE for the UPP 00Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-UPP test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-UPP test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.
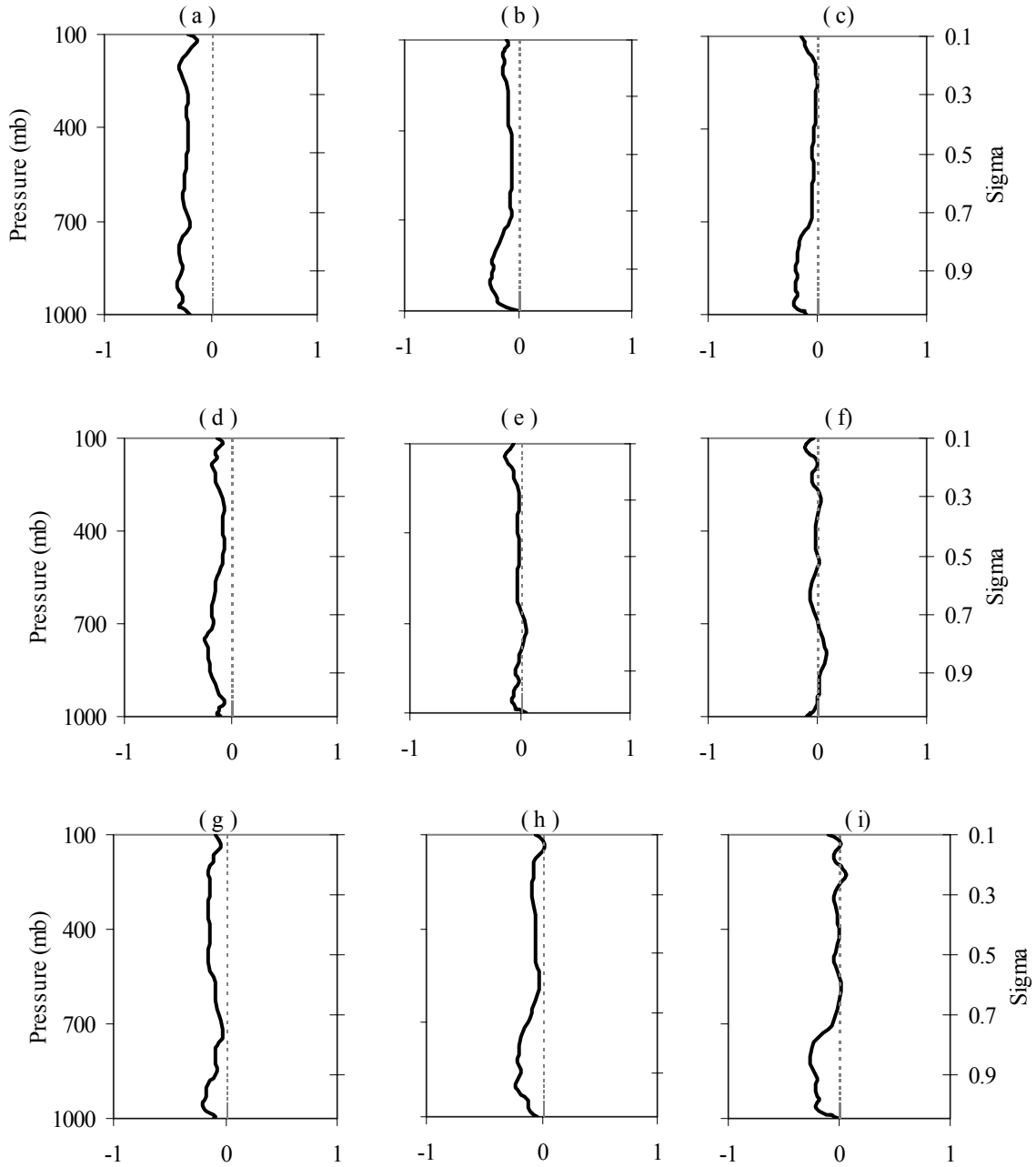
81
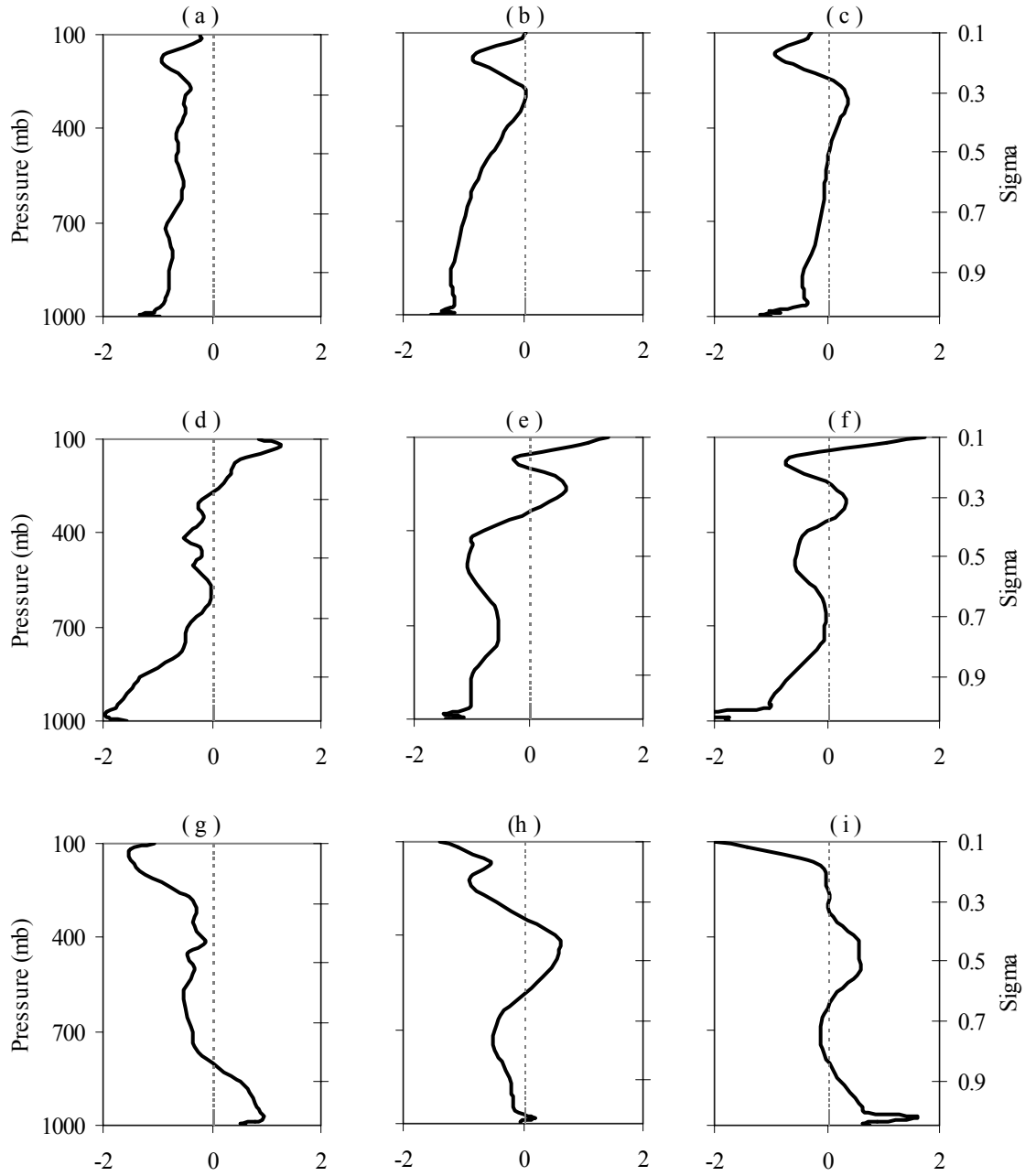
Figure 45. Average differences (C1-C2) in the geopotential height RMSE for the SFC 00Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-SFC test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-SFC test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.

82
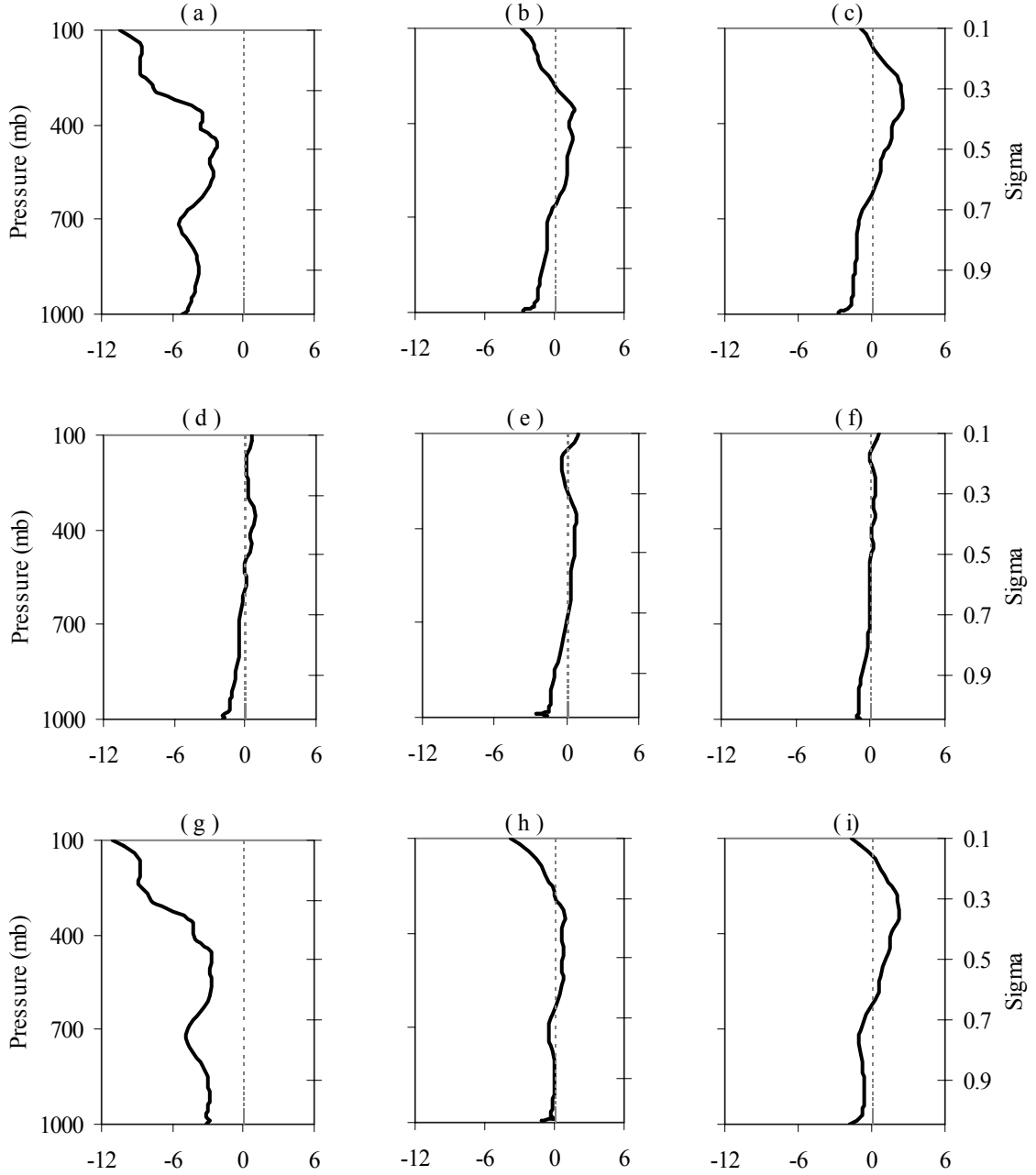
Figure 46. Average differences (C1-C2) in the geopotential height RMSE for the TOT 12Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-TOT test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-TOT test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.
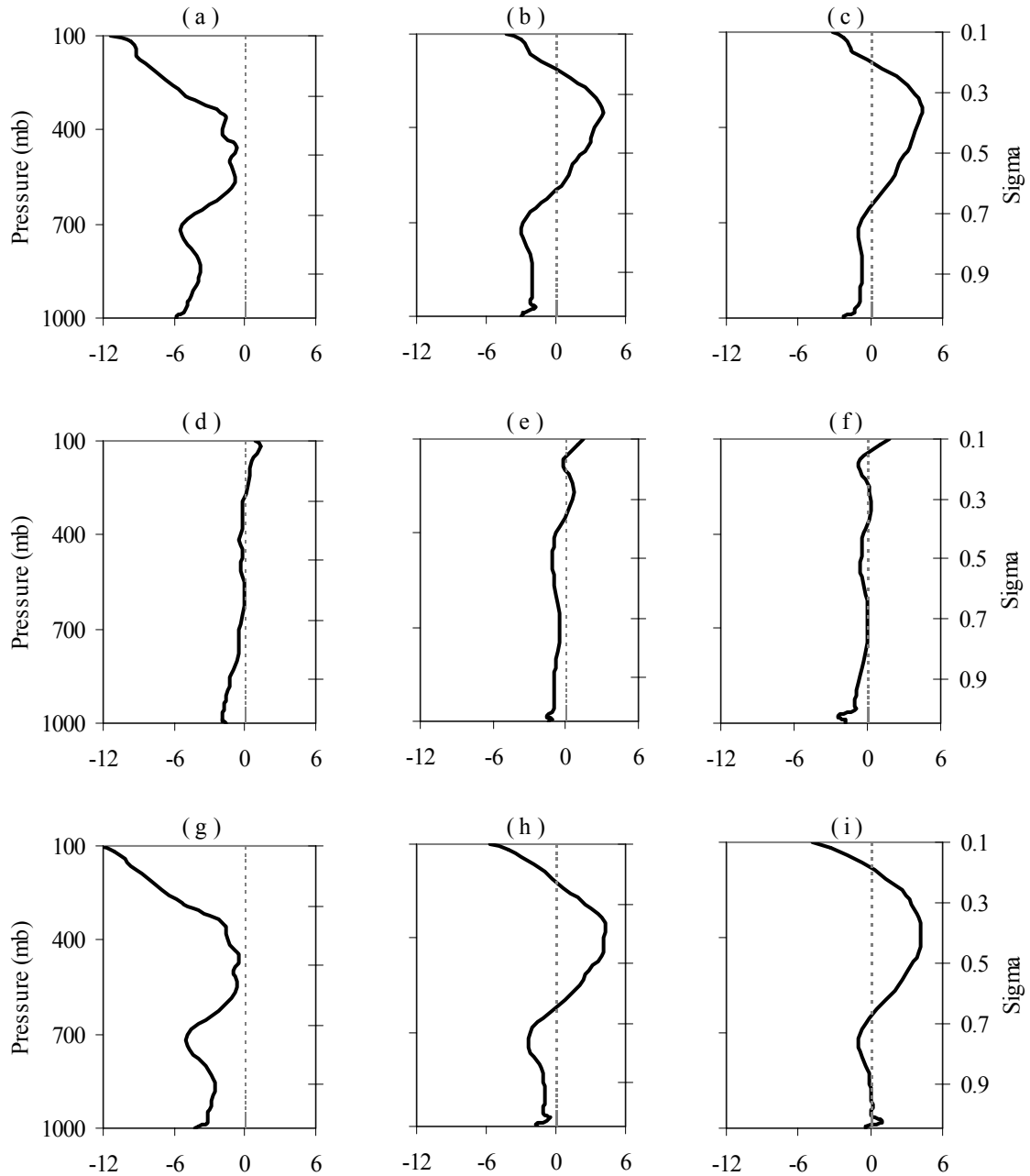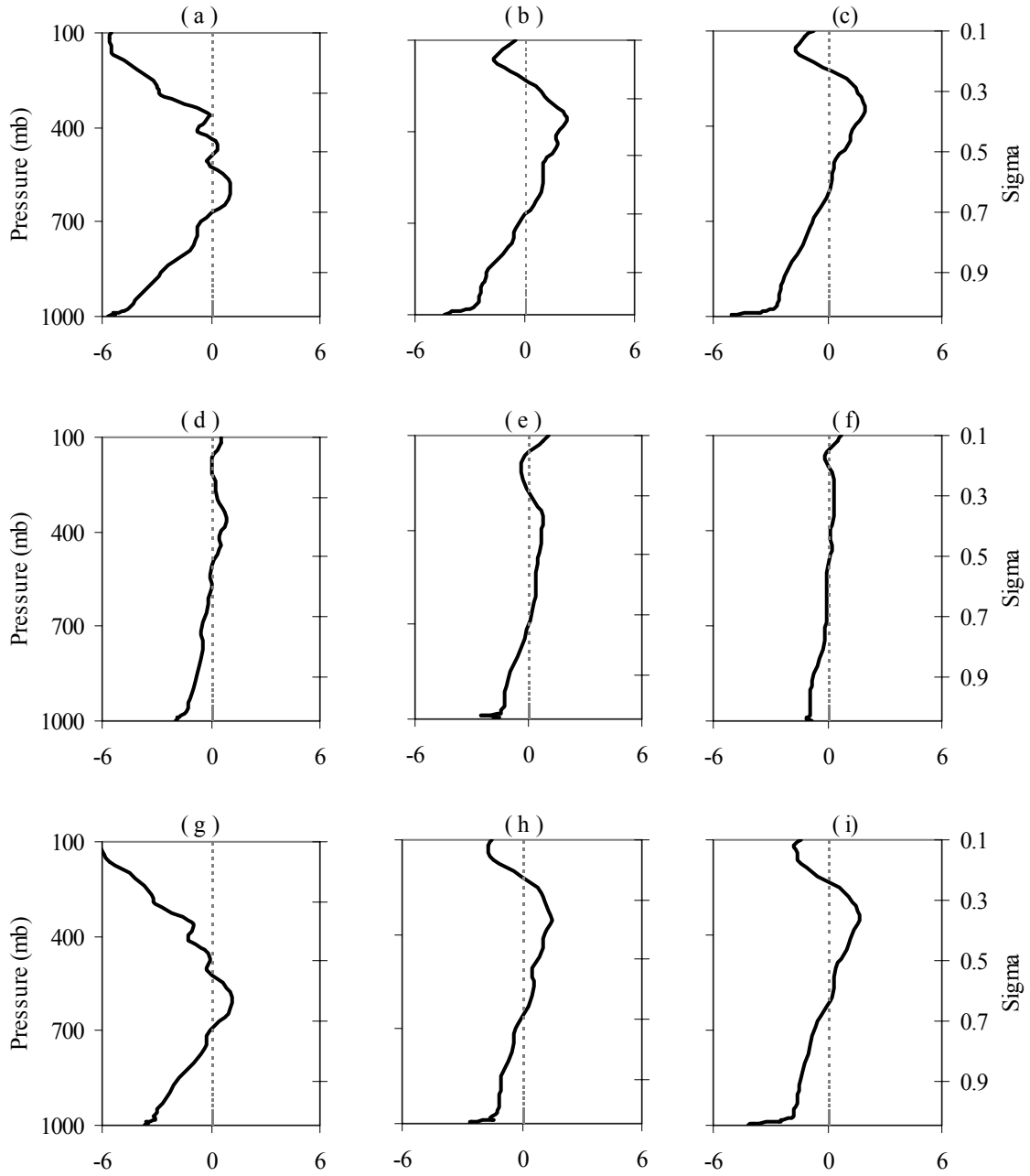
Figure 47. Average differences (C1-C2) in the geopotential height RMSE for the UPP 12Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-UPP test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-UPP test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.
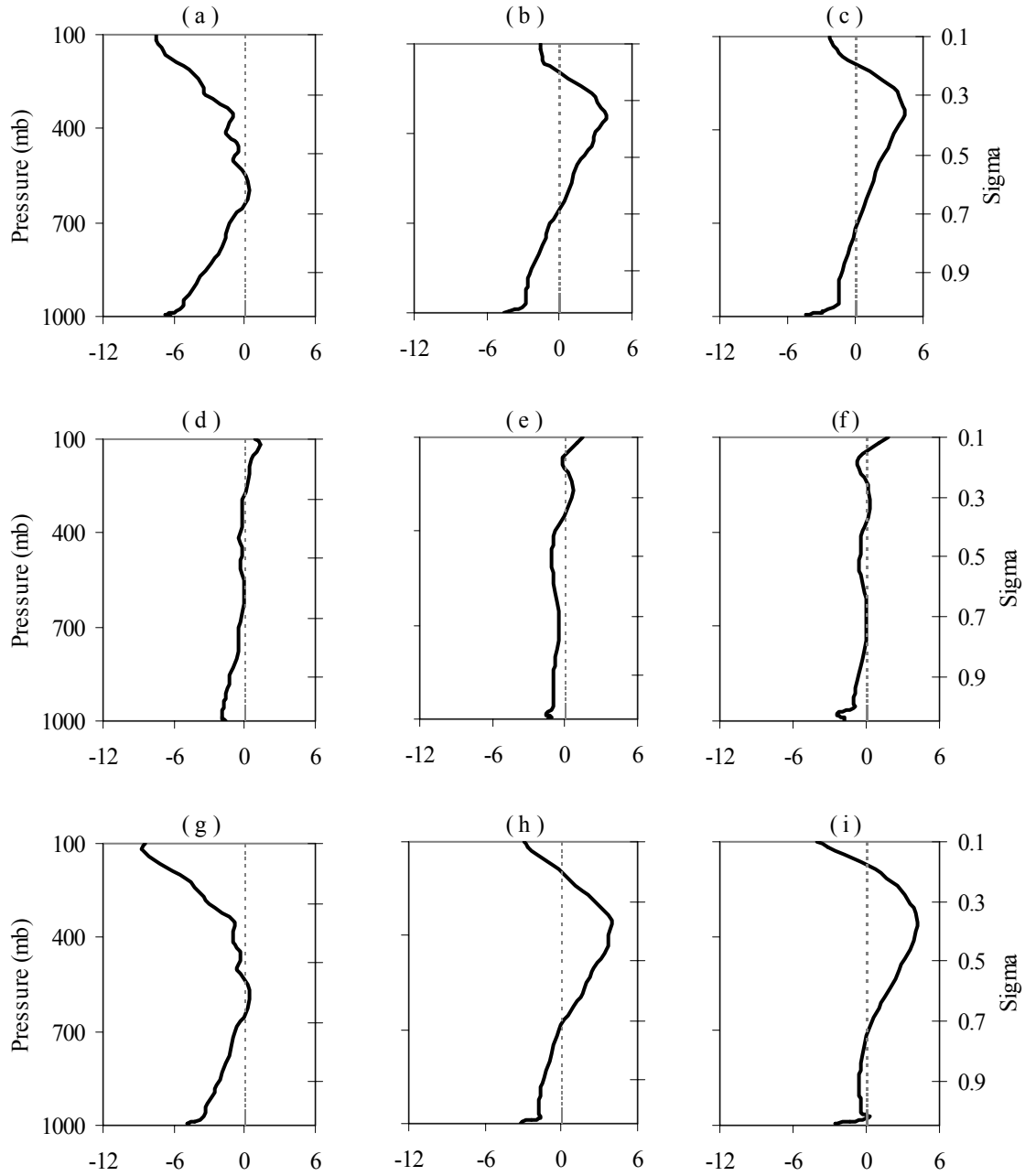
Figure 48. Average differences (C1-C2) in the geopotential height RMSE for the SFC 12Z initialization case. Charts (a), (b), and (c) show the RMSE differences of the C1-SFC test case for the 0, 12, and 24-hour forecasts respectively. Charts (d), (e), and (f) show the RMSE differences of the C2-SFC test case for the 0, 12, and 24-hour forecasts. Charts (g), (h), and (i) show the differences of (C1-C2) for the 0, 12, and 24-hour forecasts.
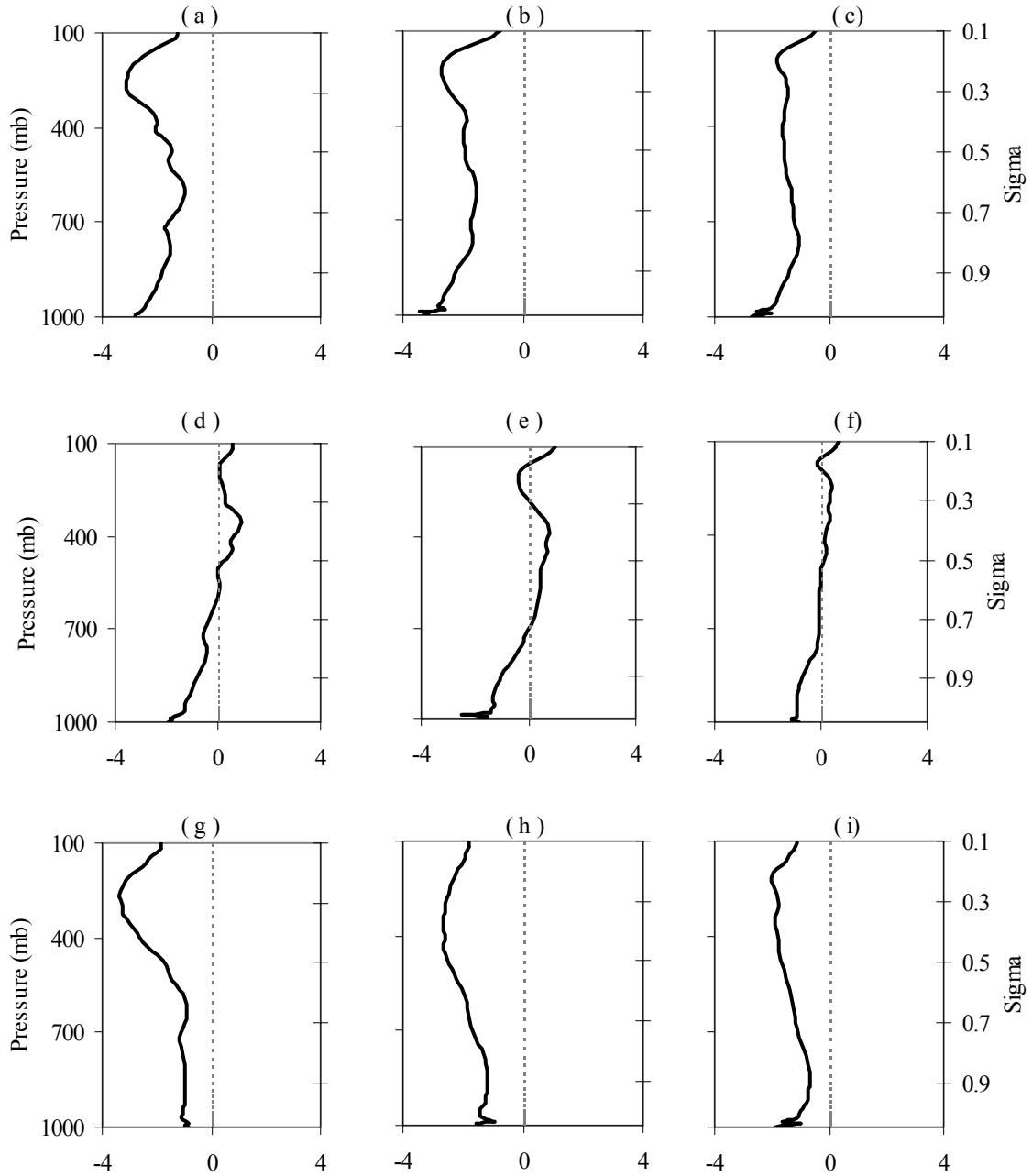
In the case of the upper air analyses, it is apparent that the different initial and

boundary conditions account for the majority of the difference between the C1 control run

and data denied test cases. The RMSE differences shown for all C1-C2 analyses are very similar to the C1-test case analyses reinforcing the importance of the first guess in the data assimilation process.

4.4.2. Surface Analyses. The surface analyses were conducted in a very similar manner to the upper air analyses with regard to accounting for the different initial and boundary conditions.

Figures 49 through 51 show the average temperature RMSE differences of the TOT, UPP, and SFC test cases respectively, Figures 52 through 54 show the mean sea level pressure RMSE differences of the various data denied test cases, and Figures 55 through 57 depict the average RMSVE differences for surface wind speed. In each of the figures, panels (a) through (c) are the results of the 00Z initialization time where panel (a) is the C1-test case RMSE difference, panel (b) is the C2-test case RMSE difference, and panel (c) shows the difference of (a) and (b). Panels (d) through (f) show the results of the 12Z initialization time where panel (d) is the C1-test case RMSE difference, panel (e) is the C2-test case RMSE difference, and panel (f) is the difference of (d) and (e). The differences shown in panels (c) and (f) of each figure are C1-C2 and estimate the effects of the use of a different first guess forecast on the MM5 model forecast accuracy.

As previously shown in the surface analyses of section 4.3, when compared with C1 the errors created by the data denial at the surface were within the accuracy of the measurement equipment. No significant deviations from this finding were noted during the current analyses. When the estimated errors associated with the different initial and boundary conditions are accounted for, the remaining RMSE differences become smaller and virtually insignificant.

Figure 49. Average differences (C1-C2) in temperature RMSE for the TOT test case. Charts (a), (b), and (c) are for the 00Z initialization case while (d), (e), and (f) are for the 12Z initialization case. Charts (a) and (d) are the RMSE differences of C1-TOT, (b) and (e) are the RMSE differences of C2-TOT, and charts (c) and (f) show the differences of C1-C2.

Figure 50. Average differences (C1-C2) in temperature RMSE for the UPP test case. Charts (a), (b), and (c) are for the 00Z initialization case while (d), (e), and (f) are for the 12Z initialization case. Charts (a) and (d) are the RMSE differences of C1-UPP, (b) and (e) are the RMSE differences of C2-UPP, and charts (c) and (f) show the differences of C1-C2.

Figure 51. Average differences (C1-C2) in temperature RMSE for the SFC test case. Charts (a), (b), and (c) are for the 00Z initialization case while (d), (e), and (f) are for the 12Z initialization case. Charts (a) and (d) are the RMSE differences of C1-SFC, (b) and (e) are the RMSE differences of C2-SFC, and charts (c) and (f) show the differences of C1-C2.
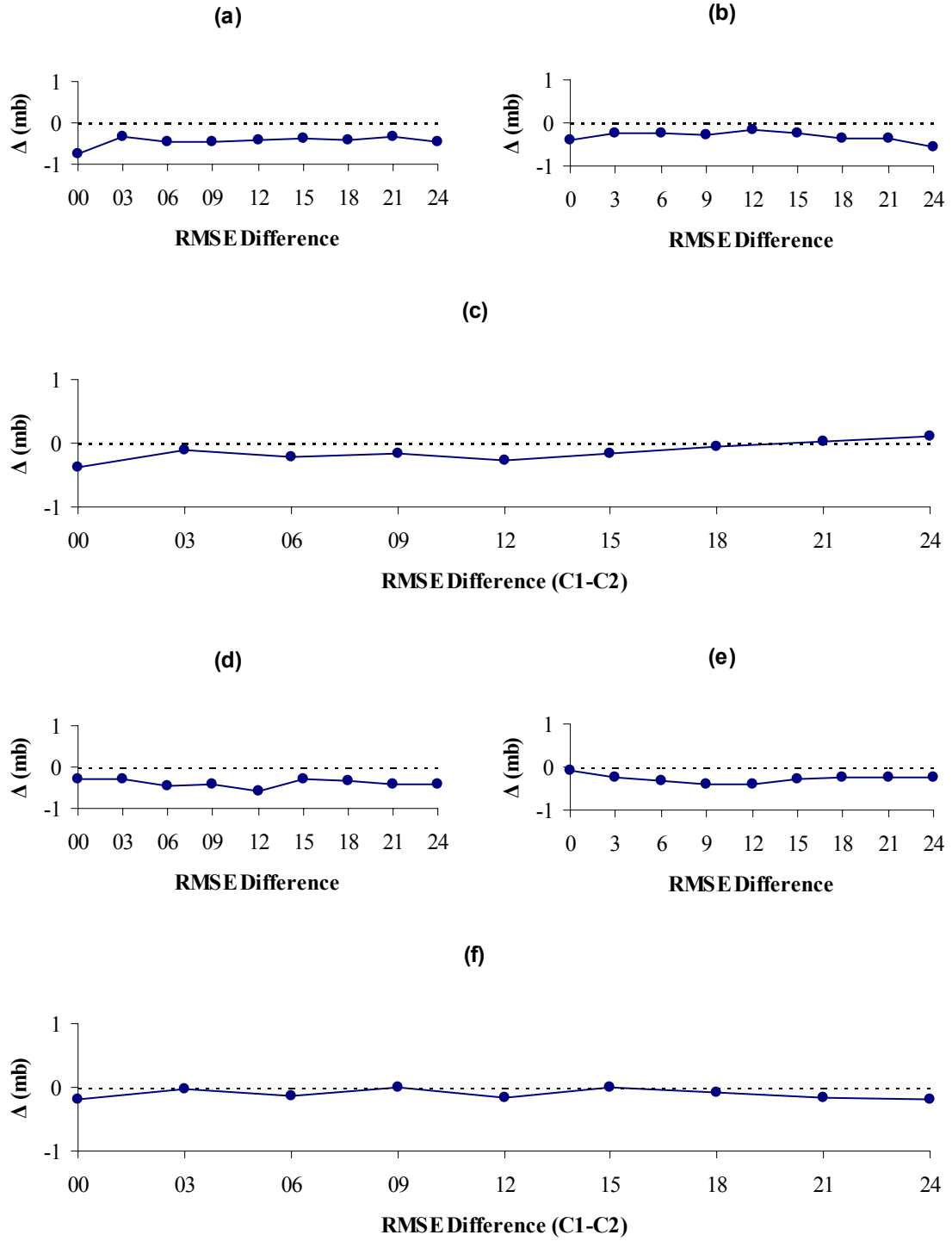
Figure 52. Average differences (C1-C2) in mean sea level pressure RMSE for the TOT test case. Charts (a), (b), and (c) are for the 00Z initialization case while (d), (e), and (f) are for the 12Z initialization case. Charts (a) and (d) are the RMSE differences of C1-TOT, (b) and (e) are the RMSE differences of C2-TOT, and charts (c) and (f) show the differences of C1-C2.
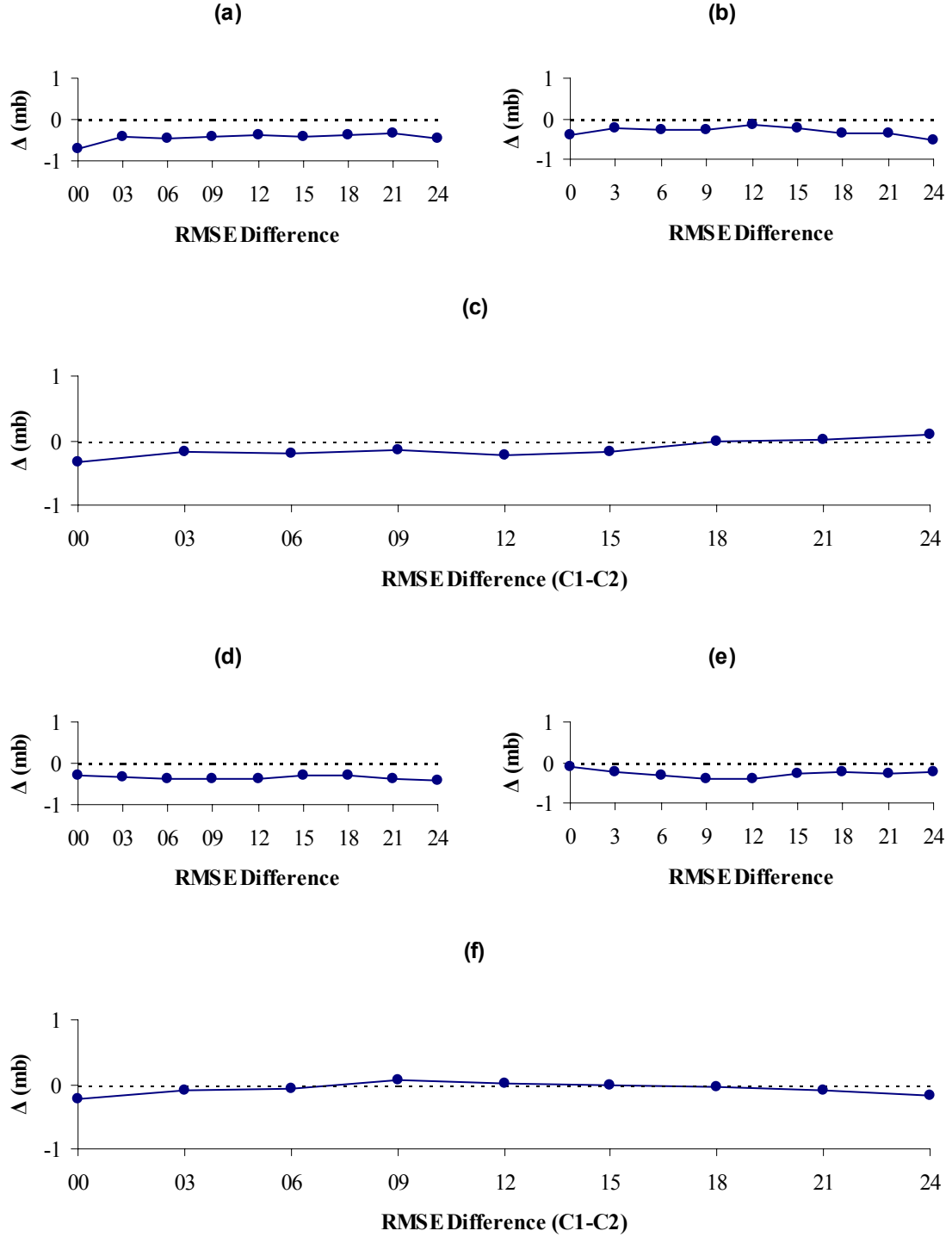
Figure 53. Average differences (C1-C2) in mean sea level pressure RMSE for the UPP test case. Charts (a), (b), and (c) are for the 00Z initialization case while (d), (e), and (f) are for the 12Z initialization case. Charts (a) and (d) are the RMSE differences of C1-UPP, (b) and (e) are the RMSE differences of C2-UPP, and charts (c) and (f) show the differences of C1-C2.
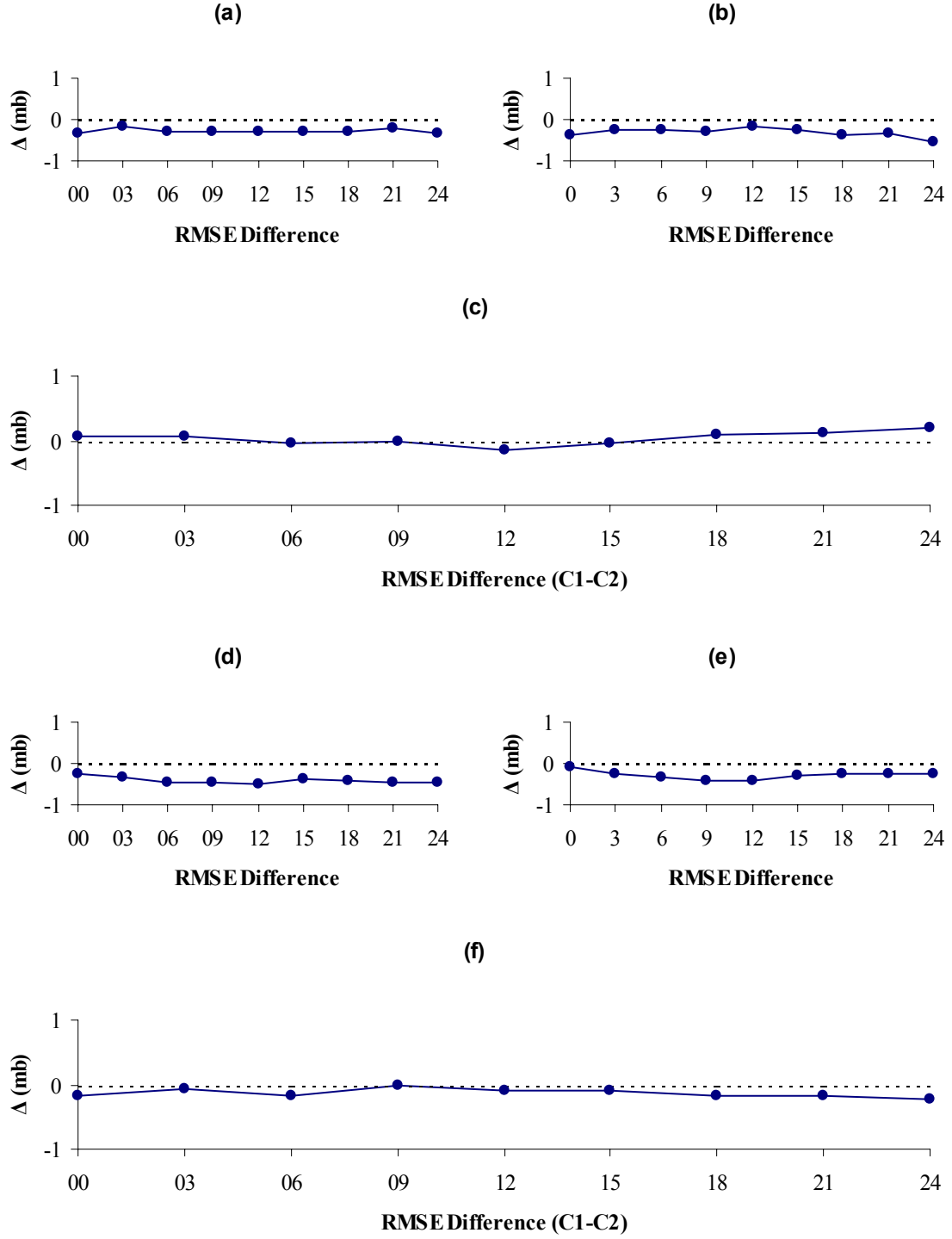
Figure 54. Average differences (C1-C2) in mean sea level pressure RMSE for the SFC test case. Charts (a), (b), and (c) are for the 00Z initialization case while (d), (e), and (f) are for the 12Z initialization case. Charts (a) and (d) are the RMSE differences of C1-SFC, (b) and (e) are the RMSE differences of C2-SFC, and charts (c) and (f) show the differences of C1-C2.
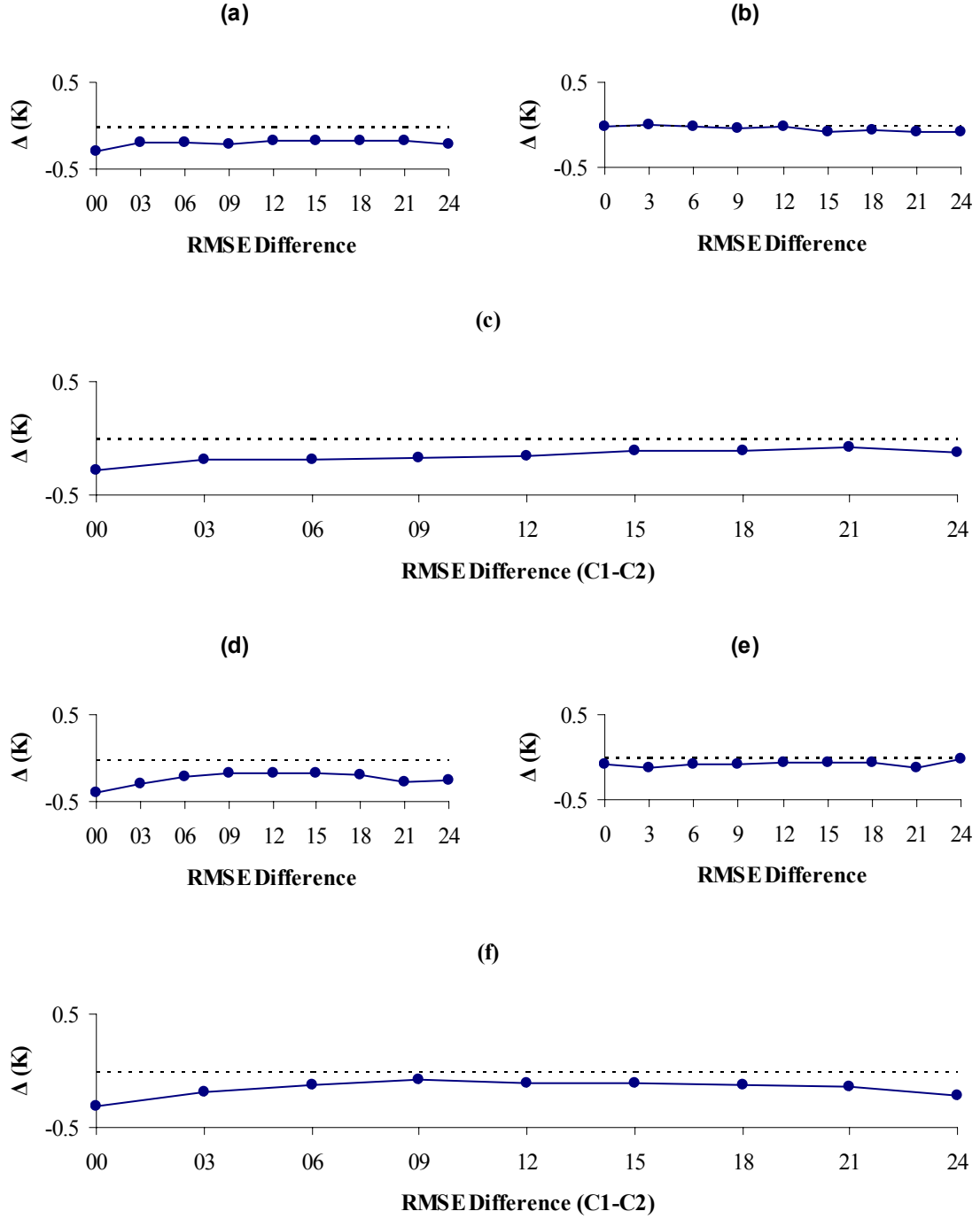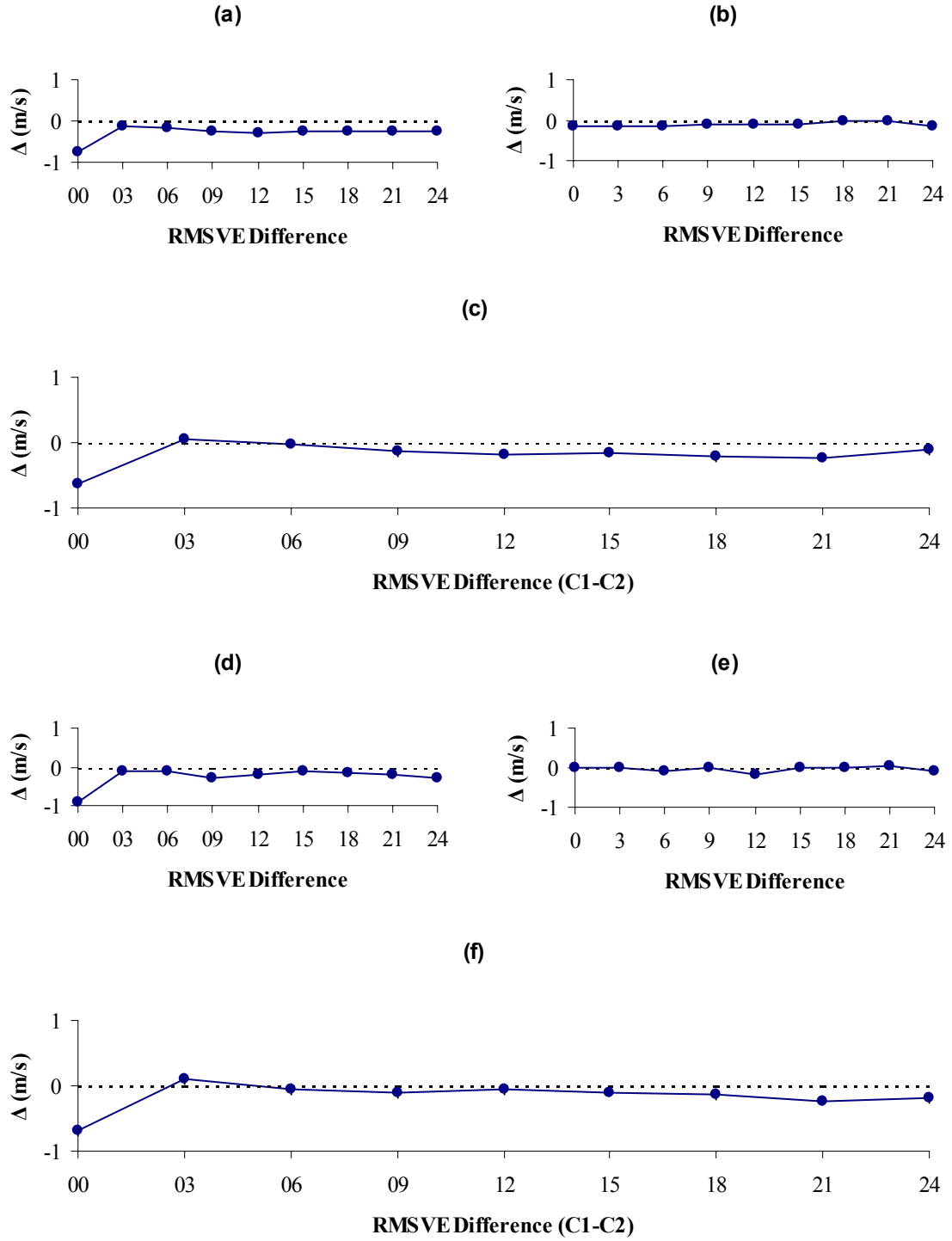
Figure 55. Average differences (C1-C2) in wind speed RMSVE for the TOT test case. Charts (a), (b), and (c) are for the 00Z initialization case while (d), (e), and (f) are for the 12Z initialization case. Charts (a) and (d) are the RMSVE differences of C1-TOT, (b) and (e) are the RMSVE differences of C2-TOT, and charts (c) and (f) show the differences of C1-C2.
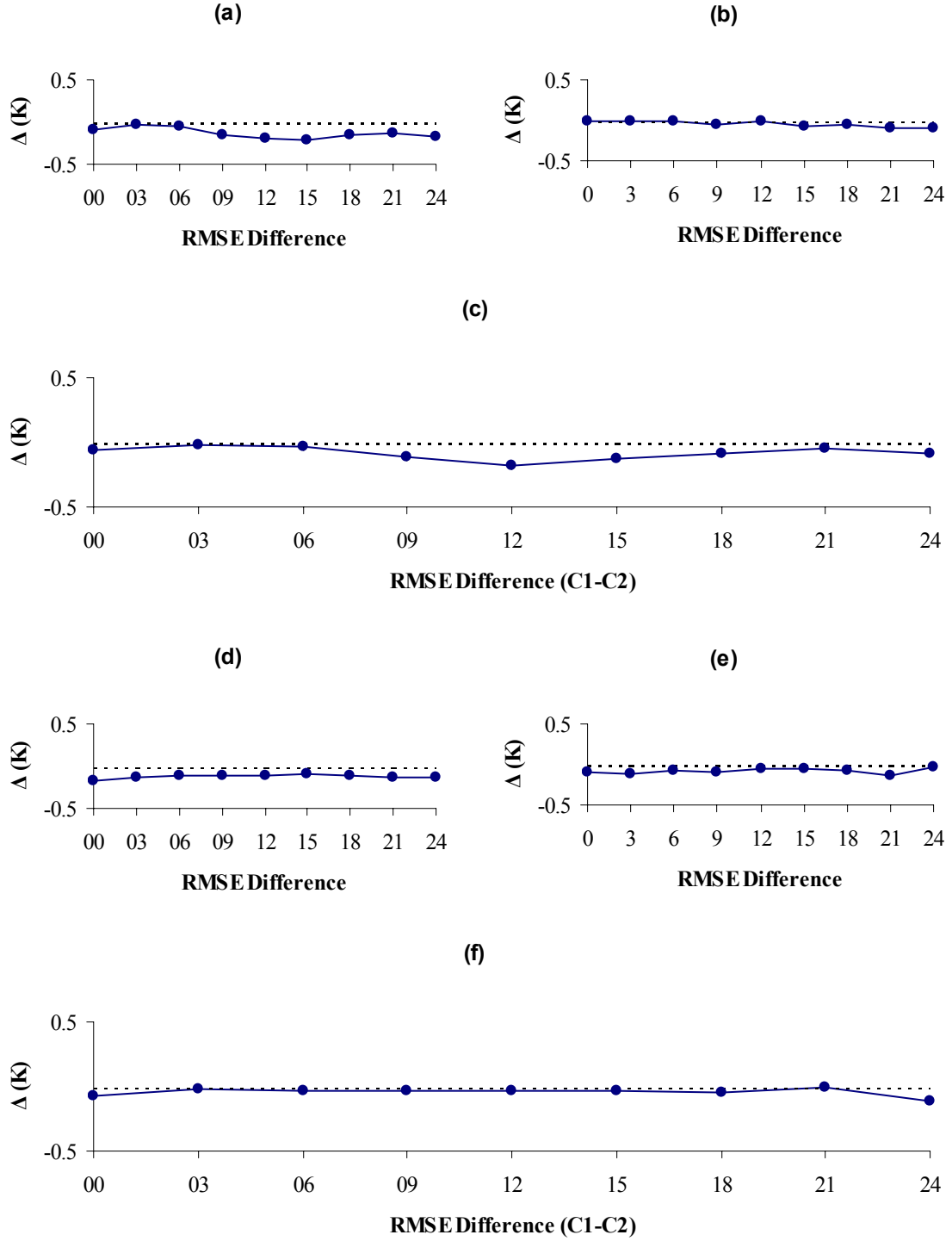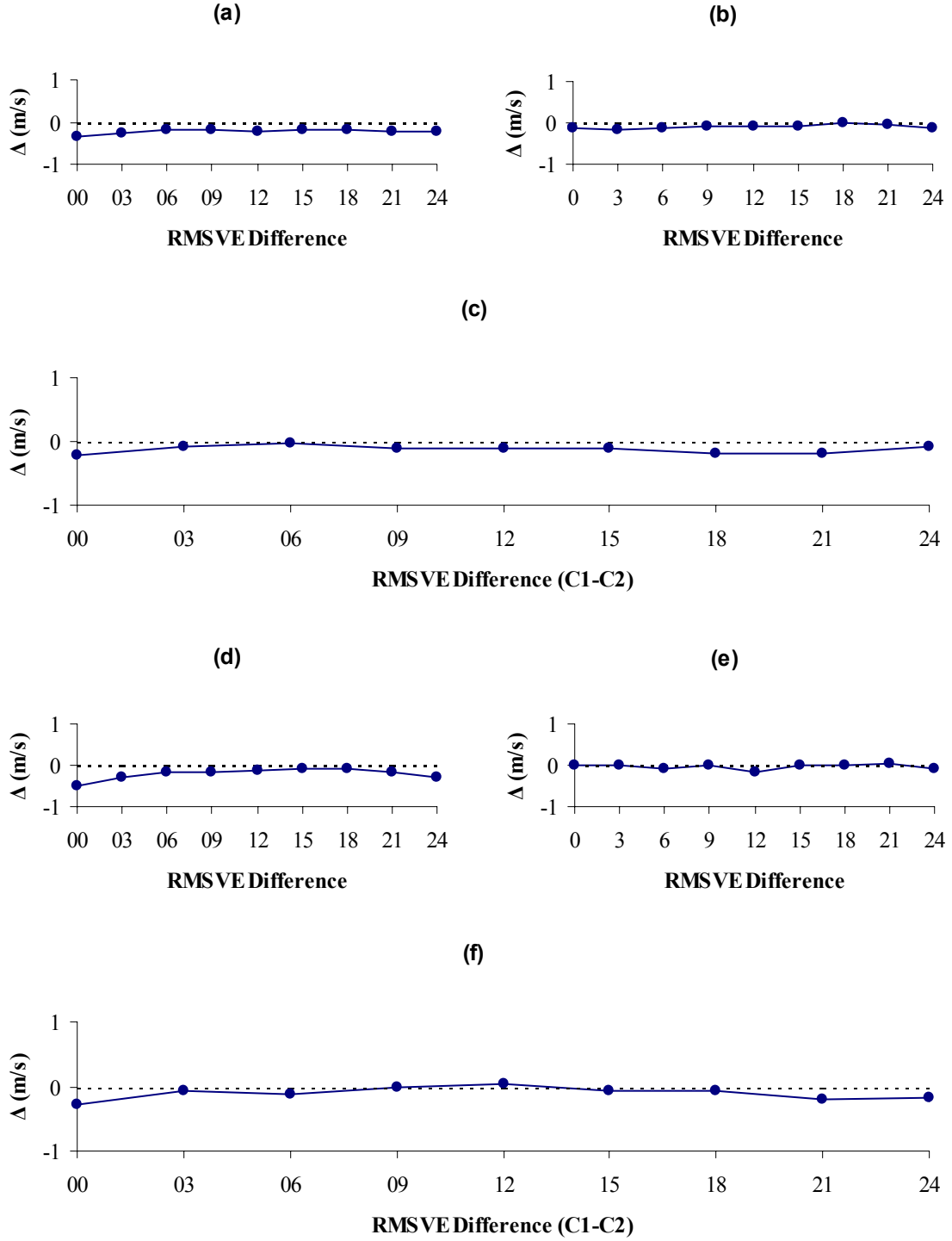
Figure 56. Average differences (C1-C2) in wind speed RMSVE for the UPP test case. Charts (a), (b), and (c) are for the 00Z initialization case while (d), (e), and (f) are for the 12Z initialization case. Charts (a) and (d) are the RMSVE differences of C1-UPP, (b) and (e) are the RMSVE differences of C2-UPP, and charts (c) and (f) show the differences of C1-C2.
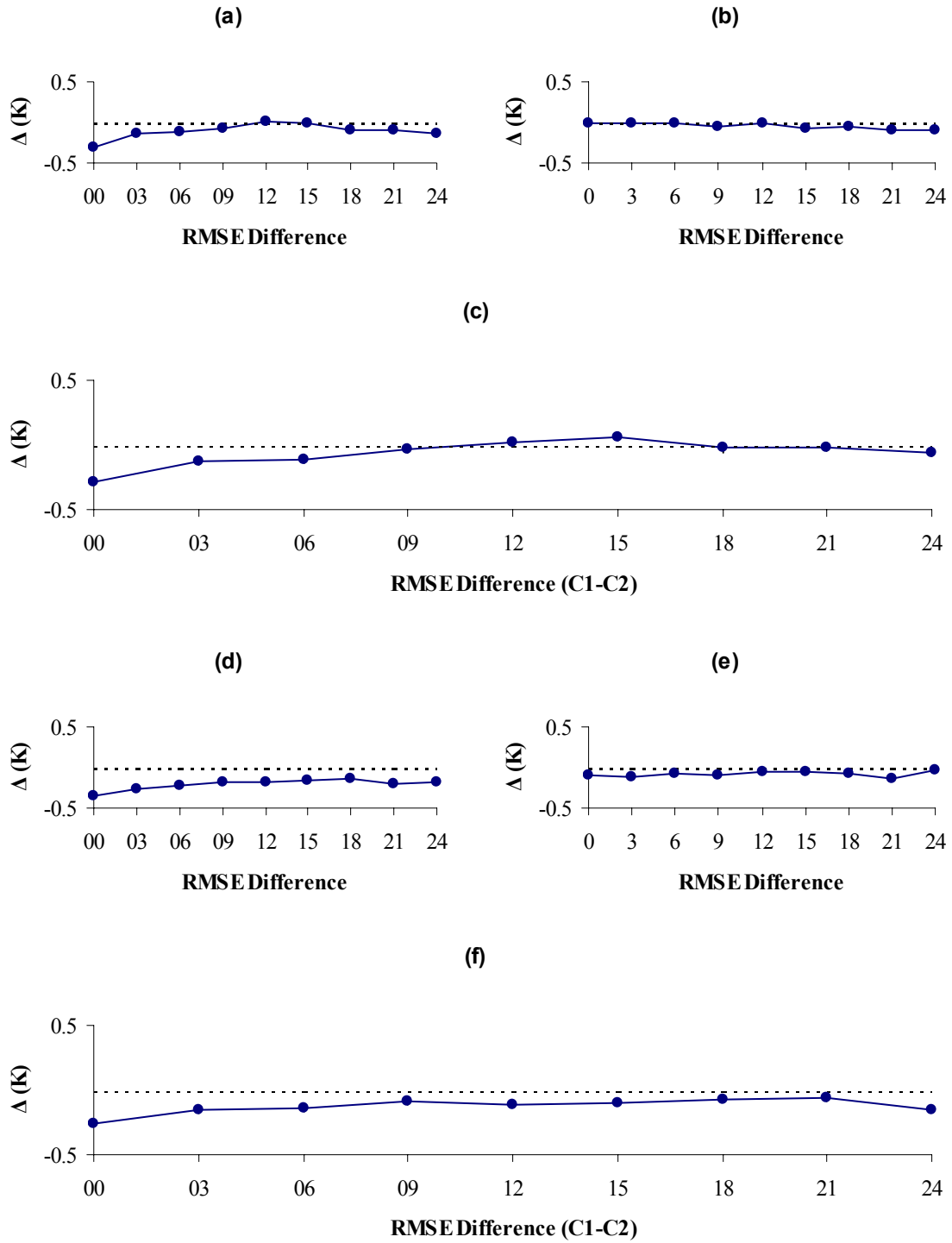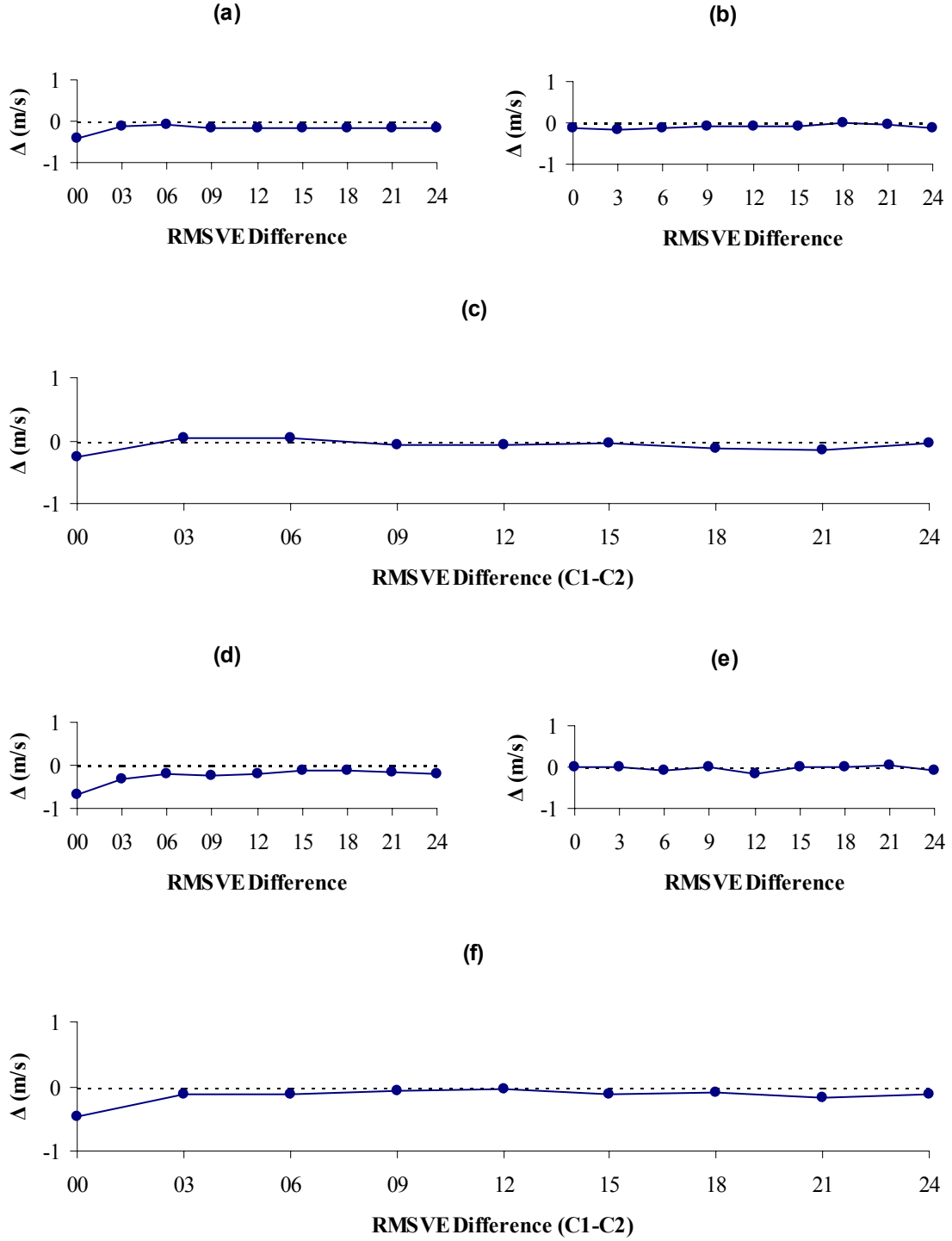
Figure 57. Average differences (C1-C2) in wind speed RMSVE for the SFC test case. Charts (a), (b), and (c) are for the 00Z initialization case while (d), (e), and (f) are for the 12Z initialization case. Charts (a) and (d) are the RMSVE differences of C1-SFC, (b) and (e) are the RMSVE differences of C2-SFC, and charts (c) and (f) show the differences of C1-C2.

V. Conclusions and Recommendations

5.1. Conclusions

The primary goal of this research was to quantify the effects of denied and limited data in MM5 model initialization on forecast accuracy. The effects of data denial and limitation on model accuracy are of critical importance to the military due to increased reliance on computer-generated weather forecasts. This study was conducted using the 3DVAR data assimilation scheme and the MM5 modeling system as configured and run operationally by the Air Force Weather Agency (AFWA). The European theater (T-03 window) was used with a 45 km horizontal grid resolution.

The input data were systematically denied in three different categories: total denial of all observations, denial of upper air observations (RAOB only), and denial of all surface observations. Two control groups for comparison were also run using all available observations as received by AFWA for the same dates and initialization times as the data denied test cases. The main control group used a 6-hour old AVN forecast which is the primary first guess forecast used by AFWA. The second control group was run using a 30-hour old AVN forecast to provide the initial and boundary conditions, which were the same data used to provide the first guess field for all of the data denied test cases. Two separate analyses were then made using the different control groups as a baseline for comparison.

The forecasts from the data denied test cases were individually compared to the two different control groups. The first set of analyses was a comparison of the main control group to the three different data denied test cases. This test configuration was chosen in an effort to replicate a real world data denial or limitation scenario in which the observations

96

would not be available for assimilation into any NWP model, including the AVN, which is used to initialize the MM5 at AFWA. The second test conducted used the secondary control group as a baseline compared to the data denied test cases. The second test was designed to approximate the error associated with the use of a different first guess forecast in the model initialization between the primary control group and the data denied test cases.

The upper air parameters analyzed for this research were geopotential heights and temperature, whereas the surface parameters were mean sea level pressure, temperature, and winds. The analyses made throughout the course of this study indicate that there are statistically significant differences between the main control group (C1) and the data denied test cases in almost every instance. The comparisons provide little evidence of meteorologically significant differences in the surface forecasts. This finding is based on the fact that the differences of the errors are well within the accuracy limitations of the equipment used to make the observations. In contrast, the upper air analyses show significant differences both statistically and meteorologically. In the upper air analyses, the largest differences emerged at the model initialization time in the cases of total data denial and upper air data denial. The magnitude of the differences rapidly diminishes as the solution of the test cases converge toward that of the control group throughout the forecast period.

In the cases of both the upper air and surface analyses, it is apparent that the different initial and boundary conditions account for the majority of the error between the C1 control run and data denied test cases. The RMSE differences shown for all C1-C2 analyses are very similar to the C1-test case analyses reinforcing the importance of the first guess in the data assimilation process. It must be noted, however, that this is an important result indicating the

significance of the observations on the data assimilation process as well. Observations that are unavailable for assimilation into the MM5 will not be available for assimilation into the AVN or other NWP models, thus forcing the use of older first guess forecasts for model initialization and leading to further model forecast degradation.

The results of the analyses show how the errors would grow if AFWA is forced to use older AVN forecasts due to data denial or limitation. The errors become much greater when the 30-hour old AVN forecast, which was made using all available observations for that particular initialization time, is used to initialize the model.

5.2. Future Research Recommendations

More in-depth model comparisons are required in order to quantify the effects of data denial. Future research should include forecasts extending beyond the 24-hour forecast period used in this study and also incorporate all seasons in order to ensure a more representative sample.

Based on the findings of this study, there are several recommendations for future research:

- Derive a method to limit the same observations to the first guess field and the test case. This will eliminate the effects on the model output of undesired observations being assimilated into the model initialization and eliminate the need for an additional control group to account for the errors associated with using different initial and boundary conditions.

- Use smaller domains with finer grid resolutions. This study used the outer domain of the European window, which includes a large amount of water along with numerous

different climatic regimes. Smaller domains would allow for more focused
concentration of observation sites without increasing the number of verification
stations.

- Perform analyses by climatic regime. Analyses of this sort would serve to decrease
  the negative effects of averaging the forecast biases.

- Compare different windows. A comparison of this sort would allow AFWA to
  determine if the MM5 is more susceptible to errors caused by the effects of data
  denial and limitation in different regions of the world.

Bibliography

British Atmospheric Data Centre (BADC), 2002: *BADC Help File: Radiosonde Data.*
    Available online at www.badc.rl.ac.uk/data/radiosonde/radhelp.html#type.

Cooperative Program for Operational Meteorology Education and Training (COMET), 2001:
    Understanding Data Assimilation: How Models Create Their Initial Conditions,
    University Corporation for Atmospheric Research. Available online at
    meted.ucar.edu/nwp/pcu1/ic6/index.htm.

Department of the Air Force, 1995: *Wind Measuring Set, AN/GMQ-20.* T.O. 31M5-
    2GMQ20-2. McClellan AFB, CA: SM-ALC/LICDA.

-----, 1997: *Digital Altimeter-Barometer, ML-658/GM.* T.O. 31M1-2GM-61. McClellan
    AFB, CA: SM-ALC/LICDA.

-----, 1998: *Ambient Temperature and Dewpoint Measuring Set, AN/FMQ-8.* T.O. 31M1-
    2FMQ8-1. McClellan AFB, CA: SM-ALC/LICDA.

Devore, J.L., 2000: *Probability and Statistics for Engineering and the Sciences* (5th Edition).
    Duxbury, 775 pp.

Dudhia, J., D. Gill, Y. Guo, K. Manning, J. Michalakes, A. Bourgeois, W. Wang, J. Wilson,
    2001: *PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's
    Guide: MM5 Modeling System Version 3*, Mesoscale and Microscale Meteorology
    Division, National Center for Atmospheric Research, 329 pp.

Grell, G.A., J. Dudhia, and D.R. Stauffer, 1994: *A Description of the Fifth Generation Penn
    State/NCAR Mesoscale Model (MM5), NCAR Tech Note 398*, National Center for
    Atmospheric Research, 117 pp.

Manning, K.W. and C.A. Davis, 1997: Verification and Senstivity Experiments for the
    WISP94 MM5 Forecasts. *Weather and Forecasting*, **12**, 719-735.

National Center for Atmospheric Research (NCAR), 2001: Available online at www.mmm.ucar.edu/mm5/3dvar.

National Oceanic & Atmospheric Administration (NOAA), 2001: Available online at www.crh.noaa.gov/mqt/fitzgerald/nwp.htm.

Pielke, R. A., 1984: *Mesoscale Meteorological Modeling.* Academic Press, 612 pp.

Ritz, R. L., 2001: Air Force Weather Agency, Offutt AFB, NE. Personal Communication.

Ritz, R. L., McAtee, M. D., Swanson, R. T., Jr., 2001: *Data Assimilation at the Air Force Weather Agency*, Preprint Volume, 14th Conference on Numerical Weather Prediction, 31 Jul - 2 Aug 2001.

Stenger, Robert A., 2000: *Sensitivity Studies on a Limited Area Mesoscale Model: An Examination of Lateral Boundary Placement, Grid Resolution and Nesting Type*, MS thesis, AFIT/GM/ENP/00M-13. School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, 193 pp.

Vaisala, 2002: *Technical Information on Radiosonde, model RS80,* Product information. Available online at http://www.vaisala.com.

Wilkes, D.S., 1998: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Wegiel, J. W., 2001: Air Force Weather Agency, Offutt AFB, NE. Personal Communication.

Appendix A. Verification Stations

Table 6. Station Verification List. All stations provided both surface and upper air observations unless otherwise noted.

STATION	ICAO	BSN	LAT	LON	Elev (m)	SFC only	Excluded
ANDOYA, Norway	ENAN	10100	69.3	16.15	14	X	
BONO, Norway	ENBO	11520	67.27	14.37	13		
ORLAND-III, Norway	ENOL	12410	63.69	9.6	7		
OSLO, Norway	ENGM	13840	60.19	11.08	204	X	
EKOFISK, Norway	-	14000	56.53	3.21	46		
STAVENGER, Norway	ENZV	14150	58.88	5.62	9		
SODANKYLA, Finland	EFSO	28360	67.36	26.65	179		
JYVASKYLA, Finland	EFJY	29350	62.4	25.68	145		
STORNOWAY, United Kingdom	EGPO	30260	58.21	-6.31	13		
KINLOSS-RAF, United Kingdom	EGQK	30660	57.65	-3.56	7	X	
BOULMER, United Kingdom	EGQM	32400	55.42	-1.59	23		
LEEMING-RAF, United Kingdom	EGXE	32570	54.3	-1.53	40	X	
WADDINGTON-RAF, United Kingdom	EGXW	33770	53.17	-0.52	70	X	
SHAWBURY-RAF, United Kingdom	EGOS	34140	52.8	-2.66	76	X	
ABERPORTH, United Kingdom	EGUC	35020	52.13	-4.56	134		
WATISHAM-RAF, United Kingdom	EGUW	35900	52.11	0.96	87	X	
BRIZE-NORTON-RAF, United Kingdom	EGVN	36490	51.75	-1.57	88	X	
HERSTOMONCEUX, United Kingdom	-	38820	50.9	0.32	54		
VANENTIA-OBSER, Ireland	-	39530	51.92	-10.25	14		
KEFLAVIK, Iceland	BIKF	40180	63.96	-22.59	52		
FAEROE-ISLANDS, Denmark	-	60110	62.01	-6.76	39		
COPENHAGEN, Denmark	-	61810	55.76	12.52	40		
DE-BILT, Netherlands	EHDB	62600	52.09	5.18	4		
KOKSIJDE, Belgium	EBFN	64000	51.07	2.65	9		
ST-HUBERT, Belgium	EBSU	64760	50.03	5.39	557		
ELSENBORN, Belgium	EBLB	64960	50.46	6.18	570		
PAYERNE, Switzerland	LSMP	66100	46.82	6.94	491		X
BREST, France	LFRB	71100	48.44	-4.42	103		
NANCY, France	LFSN	71800	48.67	6.22	217		
LYON, France	LFLL	74810	45.73	5.07	240		
BORDEAUX, France	LFBD	75100	44.82	-0.7	61		
NIMES, France	LFME	76450	43.86	4.39	62		
NICE, France	LFMN	76900	43.65	7.19	10	X	X
AJACCIO, France	LFKJ	77610	41.92	8.79	9		
ZARAGOZA, Spain	LEZG	81600	41.67	-1.01	258		
MADRID, Spain	LEMD	82210	40.44	-3.55	582		

Table 6. (continued).

STATION	ICAO	BSN	LAT	LON	Elev (m)	SFC only	Excluded
GIBRALTAR	LXGB	84950	36.15	-5.35	5		
EMDEN, Germany	EDWE	102000	53.38	7.23	1		
BERGEN, Germany	ETGB	102380	52.82	9.92	69		
BERLIN/TEMPLEHOF, Germany	EDDI	103840	52.46	13.39	49	X	
ESSEN, Germany	EDZE	104100	51.4	6.97	161		
IDAR-OBERSTEIN, Germany	ETGI	106180	49.69	7.32	377		
KUEMMERSBRUK, Germany	ETGK	107710	49.42	11.89	418		
LINZ, Austria	LOWL	110100	48.23	14.2	313		
INNSBRUK, Austria	LOWI	111200	47.26	11.35	581		X
GRAZ, Austria	LOWG	112400	47	15.42	347		X
PRAGUE, Czechoslovakia	-	115200	50	14.45	304		
LEBA, Poland	-	121200	54.75	17.52	2		
POZNAN, Poland	EPPO	123300	52.42	16.82	92	X	
SZEGED, Hungary	LHUD	129820	46.25	20.09	83		
LUUBLUANA, Slovakia	-	140150	46.07	14.51	316		
KRIZEVCI, Croatia	-	142400	45.82	16.02	123		
CLUJ-NAPOCA, Romania	LRCL	151200	46.78	23.56	413		
BUCHAREST, Romania	LRBS	154200	44.5	26.13	91		
SOFIA, Bulgaria	LBSF	156140	42.65	23.38	595		
MILAN, Italy	LIML	160800	45.42	9.27	103		
PRATICA-DI-MARE, Italy	LIRE	162450	41.65	12.42	12		
BRINDISI, Italy	LIBR	163200	40.65	17.95	10		
TRAPANI, Italy	LICT	164290	37.92	12.5	7		
CAGLIARI, Italy	LIEE	165600	39.25	9.04	1		
THESSAL-ONIKI, Greece	LGTS	166220	40.51	22.97	4		
ATHENS, Greece	LGAT	167160	37.9	23.73	15		
IRAKLION, Greece	LGIR	167540	35.32	25.18	39		
ESKISEHIR, Turkey	LTBI	171240	39.78	30.56	785	X	
AKROTIRI, Cypris	LCRA	176010	34.57	32.98	23	X	
NICOSIA, Cypris	LCNC	176070	35.15	33.4	161		
ARHANGELSK, Russia	ULAA	225500	64.52	40.46	13		
PETROZAVODSK, Russia	-	228200	61.82	34.26	112	X	
KARGOPOL, Russia	-	228450	61.5	38.92	121		
SYKTYVKAR, Russia	UUYY	238040	61.71	50.82	119		
TALLIN, Russia	EETN	260380	59.34	24.79	44		
RIGA, Russia	UMRR	264220	56.96	24.04	7		

Table 6. (continued).

STATION	ICAO	BSN	LAT	LON	Elev (m)	SFC only	Excluded
KAUNAS, Russia	-	266290	54.88	23.88	75	X	
KALININGRAD, Russia	UMKK	267020	54.69	20.61	27	X	
VOLOGDA, Russia	ULWW	270370	59.23	39.86	131		
PENZA, Russia	UWPP	279620	53.13	45.01	174		
GOMEL, Russia	UMGG	330410	52.44	31	127	X	
KISINEV, Russia	UKII	338150	47.01	28.86	180	X	
SARATOV, Russia	UWSS	341720	51.57	46.03	156	X	
ROSTOV-NA-DONU, Russia	URRR	347310	47.25	39.82	77		
DAMASCUS, Syria	OSDI	400800	33.42	36.51	609	X	
MAFRAQ, Jordan	OJMF	402650	32.36	36.25	687		
MEKNES, Morocco	GMFM	601500	33.88	-5.52	549	X	
DAR-EL-BEIDA, Algeria	DAAG	603900	36.71	3.25	25		
BECHAR, Algeria	DAOR	605710	31.61	-2.22	773		
TUNIS, Tunisia	DTTA	607150	36.82	10.23	4		
TOZEUR, Tunisia	DTTZ	607600	33.92	8.1	93	X	
BENINA, Libya	HLLB	620530	32.07	20.26	132		
MERSA-MATRUH, Egypt	HEMM	623060	31.32	27.22	30		
EL-ARISH, Egypt	HEAR	623370	31.07	33.82	32		

Vita

First Lieutenant Robert Wayne Evans graduated from Nurnberg American High School in June, 1986. He subsequently enlisted in the Air Force in December of the same year as a Precision Measurement Equipment Laboratory (PMEL) Technician.

He graduated from technical training at Lowry AFB, Colorado in September 1987. His first operational assignment was to Williams AFB, Arizona where he remained until May 1990. He was then assigned to McGuire AFB, New Jersey until August 1992. His next assignment was to San Vito Air Station, Italy where he served as the PMEL Lab Chief. In February 1994 he was assigned to the Aeronautics Laboratory at the United States Air Force Academy and worked as an Engineering Technician.

In May 1996 he was selected for the Airman Education and Commissioning Program. He was transferred to the University of Nebraska, Lincoln as a student in December 1996. He graduated in August 1998 with a Bachelor of Science degree in Meteorology/Climatology. He immediately went to Officer Training School at Maxwell AFB, Alabama where he graduated and was commissioned on 13 November 1998. His first assignment in the weather career field was with the 92nd Operations Support Squadron at Fairchild AFB, Washington as the Wing Weather Officer. In August 2000, he began the Meteorology program at the Air Force Institute of Technology. Upon graduation, he will be assigned to the Air Force Weather Agency, Offutt AFB, Nebraska.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
26-03-2002	Master's Thesis	Jun 2001 - Mar 2002

4. TITLE AND SUBTITLE
QUANTIFICATION OF THE EFFECTS OF DATA DENIAL AND LIMITATION IN MM5 INITIALIZATION ON FORECAST ACCURACY

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Evans, Robert W., First Lieutenant, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
Graduate School of Engineering and Technology (AFIT/EN)
2950 P. Street, Building 640
WPAFB, OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GM/ENP/02M-03

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFWA/DNXM
ATTN: Dr. Jerry Wegiel
106 Peacekeeper Dr.
Offutt AFB, NE 68113-4039
DSN: 271-3893

88 WS/DO
ATTN: Maj. Peter Roohr
2049 Monahan Way, Bldg 91
Wright-Patterson AFB, OH 45433-7204
DSN: 785-2316

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Using the 3-Dimensional Variational Analysis data assimilation scheme and the MM5, input observations were denied in three different categories: total, upper air, and surface observation denial. Two control groups were run using all available data as received by AFWA. The main control group used a 6-hour old first guess as a baseline. The data denied test cases and the secondary control group used a 30-hour old first guess because it was not possible to deny data from the first guess. The secondary control group was used to estimate errors resulting from the use of different first guess forecasts between the main control group and the test cases for all forecast times. The analyses show statistically significant differences between the main control group and test cases in almost every instance. However, the surface comparisons provide little evidence of significant meteorological differences due to the relatively small magnitude of the differences in root mean square error and bias. In the upper air analyses, the largest differences emerged at the model initialization time. The magnitude of the differences rapidly diminishes as the solution of the test cases converge toward that of the main control group throughout the forecast period. Errors associated with the use of different initial and boundary conditions account for most of the resulting differences.

15. SUBJECT TERMS
Mesoscale Model, MM5, 3DVAR, 3-Dimensional Variational Analysis, Air Force Weather Agency, AFWA MM5, Numerical Weather Prediction, Model Comparison, Atmospheric Temperature, Weather

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Lt Col Michael K. Walters, Ph.D., ENP
U	U	U	UU	120	19b. TELEPHONE NUMBER (Include area code) (937) 255-3636, ext 4681